FEDERAL SUPERCOMPUTER PROGRAMS AND POLICIES

HEARING

BEFORE THE

SUBCOMMITTEE ON
ENERGY DEVELOPMENT AND APPLICATIONS

AND THE

SUBCOMMITTEE ON
SCIENCE, RESEARCH AND TECHNOLOGY

OF THE

COMMITTEE ON
SCIENCE AND TECHNOLOGY
HOUSE OF REPRESENTATIVES

NINETY-NINTH CONGRESS

FIRST SESSION

JUNE 10, 1985

[No. 44]

Printed for the use of the
Committee on Science and Technology

FEDERAL SUPERCOMPUTER PROGRAMS AND POLICIES

HEARING

BEFORE THE

SUBCOMMITTEE ON ENERGY DEVELOPMENT AND APPLICATIONS

AND THE

SUBCOMMITTEE ON SCIENCE, RESEARCH AND TECHNOLOGY

OF THE

COMMITTEE ON SCIENCE AND TECHNOLOGY HOUSE OF REPRESENTATIVES

NINETY-NINTH CONGRESS

FIRST SESSION

JUNE 10, 1985

[No. 44]

Printed for the use of the
Committee on Science and Technology

U.S. GOVERNMENT PRINTING OFFICE

53-831 O WASHINGTON : 1986

COMMITTEE ON SCIENCE AND TECHNOLOGY

DON FUQUA, Florida, *Chairman*

ROBERT A. ROE, New Jersey
GEORGE E. BROWN, Jr., California
JAMES H. SCHEUER, New York
MARILYN LLOYD, Tennessee
TIMOTHY E. WIRTH, Colorado
DOUG WALGREN, Pennsylvania
DAN GLICKMAN, Kansas
ROBERT A. YOUNG, Missouri
HAROLD L. VOLKMER, Missouri
BILL NELSON, Florida
STAN LUNDINE, New York
RALPH M. HALL, Texas
DAVE McCURDY, Oklahoma
NORMAN Y. MINETA, California
MICHAEL A. ANDREWS, Texas
BUDDY MacKAY,* Florida
TIM VALENTINE,* North Carolina
HARRY M. REID, Nevada
ROBERT G. TORRICELLI, New Jersey
FREDERICK C. BOUCHER, Virginia
TERRY BRUCE, Illinois
RICHARD H. STALLINGS, Idaho
BART GORDON, Tennessee
JAMES A. TRAFICANT, Jr., Ohio

MANUEL LUJAN, Jr.,* New Mexico
ROBERT S. WALKER, Pennsylvania
F. JAMES SENSENBRENNER, Jr., Wisconsin
CLAUDINE SCHNEIDER, Rhode Island
SHERWOOD L. BOEHLERT, New York
TOM LEWIS, Florida
DON RITTER, Pennsylvania
SID W. MORRISON, Washington
RON PACKARD, California
JAN MEYERS, Kansas
ROBERT C. SMITH, New Hampshire
PAUL B. HENRY, Michigan
HARRIS W. FAWELL, Illinois
WILLIAM W. COBEY, Jr., North Carolina
JOE BARTON, Texas
D. FRENCH SLAUGHTER, Jr., Virginia
DAVID S. MONSON, Utah

HAROLD P. HANSON, *Executive Director*
ROBERT C. KETCHAM, *General Counsel*
REGINA A. DAVIS, *Chief Clerk*
JOYCE GROSS FREIWALD, *Republican Staff Director*

SUBCOMMITTEE ON ENERGY DEVELOPMENT AND APPLICATIONS

DON FUQUA, Florida, *Chairman*

ROBERT A. ROE, New Jersey
RALPH M. HALL, Texas
FREDERICK C. BOUCHER, Virginia
TERRY BRUCE, Illinois
JAMES A. TRAFICANT, Jr., Ohio
DOUG WALGREN, Pennsylvania
ROBERT A. YOUNG, Missouri
DAVE McCURDY, Oklahoma
RICHARD H. STALLINGS, Idaho
GEORGE E. BROWN, Jr., California

F. JAMES SENSENBRENNER, Jr., Wisconsin
CLAUDINE SCHNEIDER, Rhode Island
RON PACKARD, California
HARRIS W. FAWELL, Illinois
WILLIAM W. COBEY, Jr., North Carolina
JOE BARTON, Texas
D. FRENCH SLAUGHTER, Jr., Virginia

WILLIAM T. HARVEY, Jr. *Staff Director*
HARLAN L. WATSON, *Science Consultant*
WILLIAM J. GORDON, *Republican Technical Consultant*
R. THOMAS WEIMER, *Republican Energy Coordinator*

SUBCOMMITTEE ON SCIENCE, RESEARCH AND TECHNOLOGY

DOUG WALGREN, Pennsylvania, *Chairman*

STAN LUNDINE, New York
NORMAN Y. MINETA, California
GEORGE E. BROWN, Jr. California
TIMOTHY E. WIRTH, Colorado
TERRY BRUCE, Illinois
TIM VALENTINE, North Carolina

SHERWOOD L. BOEHLERT, New York
DON RITTER, Pennsylvania
PAUL B. Henry, Michigan
WILLIAM W. COBEY, Jr., North Carolina

EZRA HEITOWIT, *Staff Director*
GRACE OSTENSO, *Science Consultant*
MARYANNE BACH, *Republican Technical Consultant*
DOUGH THOMPSON, *Special Assistant to the Republican Members*

* Ranking Republican Member.
** Serving on Committee on the Budget for 99th Congress.

CONTENTS

WITNESSES

Page

Appendix: Additional Materials for the Record

FEDERAL SUPERCOMPUTER PROGRAMS AND POLICIES

MONDAY, JUNE 10, 1985

HOUSE OF REPRESENTATIVES, COMMITTEE ON SCIENCE AND TECHNOLOGY, SUBCOMMITTEE ON ENERGY DEVELOPMENT AND APPLICATIONS, AND THE SUBCOMMITTEE ON SCIENCE, RESEARCH AND TECHNOLOGY,

Tallahassee, FL.

The subcommittee met, pursuant to call, at 9 a.m., in the Conference Center, Florida State University, Tallahassee, FL, Hon. Don Fuqua (chairman of the Subcommittee on Energy Development and Applications) presiding.

Mr. FUQUA. Today's hearing concerns the Federal supercomputer programs and policies. We are pleased to be here to precede this important national supercomputer conference.

The Committee on Science and Technology has played an important role in this Nation's supercomputer development. Basic research programs funded by the committee impact the supercomputer industry in a number of ways.

First, we strongly support the acquisition of the latest generation supercomputers by the National Science Foundation and the Department of Energy to provide the computing capacity. The establishment of the Supercomputer Computational Research Institute here at Florida State by the Department of Energy is such an example.

Second, the committee provides funding to make these machines more readily accessible to researchers through communications networks and software support. Clearly, this is a critical area where we have a long way to go, as we will hear from several of today's witnesses. However, we have made a start and we hope to do much more.

Third, the committee supports researchers who are trying to develop the next generation of supercomputers—advanced supercomputers which have the potential to be hundreds or even thousands of times more powerful than today's machines.

We are looking forward to hearing from today's expert witnesses about their present plans and future requirements for supercomputing networking, software, and research and development and what the Federal Government might do to enhance their capabilities in the development and application of this important technology.

Mr. LUJAN. Mr. Chairman, let me join you in saying how happy we on the committee are to be here in your home district for this hearing to review Federal supercomputer programs and policies.

The Federal Government has a long-standing tradition in supercomputer usage, dating back to the research programs of the Department of Energy laboratories like Los Alamos, Sandia, and Lawrence Livermore. Two of these labs are in my home State of New Mexico and I am very familiar with their work.

The Committee on Science and Technology, in its capacity of authorizing and overseeing all Federal civilian scientific research and development, has taken particular interest in promoting quality and state-of-the-art computational resources to the scientific community. This committee planted the seeds for what is now the "Advanced Scientific Computing Initiative," organized through the National Science Foundation. The goal of this initiative, of course, is to enhance the access and quality of U.S. supercomputers.

In addition, recognizing the interests of DOE, NSF, the Defense Advanced Research Program Agency [DARPA], the National Bureau of Standards [NBS], the National Aeronautics and Space Administration [NASA], and other Federal R&D agencies in supercomputer technology, this committee has been a strong promoter of coordination within the Federal Government as well as of partnerships with the U.S. industrial, scientific and engineering communities. This conference, I feel, will certainly promote that spirit.

Again, Mr. Chairman, I would like to thank you for bringing us together for this conference. I feel we will all benefit by this gathering.

Mr. FUQUA. Thank you, Mr. Lujan.

I would also like to welcome to Florida, Congressman Jim Sensenbrenner from Wisconsin, who is the ranking Republican on the Subcommittee on Energy Development and Applications that deals with—that has the responsibility of the DOE's basic energy and nuclear science programs.

Jim, we are glad to have you and welcome to Tallahassee.

Mr. SENSENBRENNER. Thank you very much, Mr. Chairman.

This is not the first time you have brought me down here. Two years ago I did attend a conference in Marianna, which I am informed is somewhat to the west of here, which was very useful in terms of investigating energy alternatives. I am excited about this conference, and I come here as someone who is more objective in that I do not have any supercomputer facilities in my district and none are being planned.

Nonetheless, supercomputers, I think, are going to be one of the critical ways that the United States can make technological advantage over other countries of the world. We have to give whatever encouragement is necessary to keep that technological edge and to develop supercomputer technology.

This committee and subcommittee have a role in providing some of the funding for that, and I am happy to participate today as well as in the future in the supercomputer activities.

Mr. FUQUA. Thank you, Jim.

Next I would like to welcome Congressman Sherwood Boehlert from New York. He is the ranking Republican member of the Science, Research and Technology Subcommittee that deals with the

budget that relates to the National Science Foundation, NBS, and some of the other areas of our overall committee jurisdiction.

Sherwood, we are glad to have you in Florida, and you were here for the rain yesterday.

Mr. BOEHLERT. I was, Mr. Chairman. I thank you very much.

I was a natural for this Committee on Science and Technology when I got to Congress 3 years ago because I had a D in high school physics.

When I joined this committee, I was stunned to hear one of my constituents, Dr. Ken Wilson, who is going to present testimony today and who is, I consider, a national asset—he is a Nobel Laureate—testify that graduate students in other countries—West Germany, Great Britain—had greater access to the latest in computer technology than did he. That concerned me, and I think it obviously concerned all of our committee.

Further investigation found that American scientists often had to travel abroad to use American-built supercomputers to complete their research.

Dr. Wilson's testimony, along with reports that Japanese supercomputers had surpassed our own, quickly made the improvement of the United States supercomputer capacity one of my top priorities, and the committee embraced this totally.

With this committee leading the way, the Federal Government has dedicated $200 million to supercomputer research and access this year. And despite a probable freeze on overall National Science Foundation spending, the budget for supercomputing is likely to go up. Even the administration proposed an increase of 11.3 percent in supercomputing funding.

As part of this effort, NSF has designated, as we all know, four supercomputer centers, one in beautiful upstate New York, which will lead the way in developing an American supercomputer network, available to students and other researchers in academia and industry.

This NSF program is a striking example of how Federal and State governments, industry and education can work together for economic and scientific progress.

I look forward to hearing our participants today and I am excited as I look at the program because I am privileged to be here among the movers and shakers of the supercomputer technology in the United States. There is a lot of room for movement and there is a need for a lot more shaking.

I am glad, Mr. Chairman, that you thought of this session today and I am just pleased to be here among all of you.

Mr. FUQUA. Thank you, Mr. Boehlert. We are glad to have you here.

I would like to insert the comments of Congressman Doug Walgren of Pennsylvania, who is the chairman of the Science, Research and Technology Subcommittee. He had hoped to be here but, regretfully, was unable to attend. He does have a prepared statement that he would like included in the record at this time.

[The prepared opening statement of Mr. Walgren follows:]

OPENING STATEMENT

DOUG WALGREN

CHAIRMAN, SCIENCE, RESEARCH AND TECHNOLOGY SUBCOMMITTEE

FEDERAL SUPERCOMPUTER PROGRAMS AND POLICIES

JUNE 10, 1985

THE PURPOSE OF OUR HEARING TODAY IS TO ASSESS THE ADEQUACY OF CURRENT FEDERAL INITIATIVES AND PLANS FOR ADDRESSING PRESENT AND FUTURE NEEDS FOR LARGE-SCALE SCIENTIFIC COMPUTING. THE FEDERAL GOVERNMENT HAS SUPPORTED THE DOMESTIC COMPUTER INDUSTRY SINCE ITS INCEPTION THROUGH RESEARCH AND DEVELOPMENT FUNDING AND PROCUREMENT. THE GOVERNMENT HAS STIMULATED THE SUPERCOMPUTER INDUSTRY BY PURCHASING OR LEASING MORE THAN 50 PERCENT OF THE PRESENT GENERATION MACHINES IN SERVICE IN THE UNITED STATES.

SEVERAL MAJOR FEDERAL EFFORTS AIMED AT STRENGTHENING U.S. SUPERCOMPUTER-RELATED CAPABILITIES HAVE RECENTLY BEEN INITIATED. THE QUESTION BEFORE US TODAY IS WHETHER OR NOT THESE PROGRAMS AND POLICIES ARE PROPERLY FOCUSED AND INTEGRATED TO MAINTAIN OUR TRADITIONAL SUPERCOMPUTER LEADERSHIP AND TO ENHANCE RESEARCH, TRAINING, AND APPLICATIONS IN SUPERCOMPUTER TECHNOLOGY.

ALTHOUGH VERY HIGH-PERFORMANCE SCIENTIFIC COMPUTERS ARE A SMALL PART OF THE TOTAL COMPUTER MARKET, THEIR STRATEGIC VALUE IN ADVANCING SCIENCE AND TECHNOLOGY IS QUITE SIGNIFICANT. IT IS IMPERATIVE THAT THE FEDERAL POLICIES THAT ACHIEVED U.S. SUPERCOMPUTER LEADERSHIP IN THE PAST BE REVITALIZED TO BE CONSISTENT WITH TODAY'S NEEDS AND BE IMPLEMENTED THROUGH DYNAMIC PROGRAMS. TO ASSURE THAT CURRENT PROGRAMS,

FUTURE TRENDS, AND EMEGING ISSUES REGARDING SUPERCOMPUTER TECHNOLOGY AND APPLICATIONS ARE APPROPRIATELY ADDRESSED, QUANTITATIVE MEASUREMENTS OF PROGRESS ARE WARRANTED. GOVERNMENT-INDUSTRY-UNIVERSITY PARTNERSHIPS WILL BE NEEDED TO DEVELOP THE TOOLS AND TECHNIQUES NECESSARY TO MEASURE PROGRESS IN TRAINING, NETWORKS, SOFTWARE, AND ARCHITECTURE FOR SUPERCOMPUTING SYSTEMS.

THE SUBCOMMITTEE ON SCIENCE, RESEARCH AND TECHNOLOGY, WHICH AUTHORIZES THE NATIONAL SCIENCE FOUNDATION (NSF), HAS ENCOURAGED THE ESTABLISHMENT OF A VIABLE ADVANCED SCIENTIFIC COMPUTING PROGRAM AT NSF. THE SUBCOMMITTEE IS PLEASED WITH THE TIMELY FASHION IN WHICH NSF DEVELOPED THE PROGRAM TO PROVIDE HIGH-QUALITY ACCESS TO ADVANCED COMPUTATIONAL FACILITIES. TO TAKE FULL ADVANTAGE OF THE POTENTIAL CONTRIBUTION OF THE NEW NATIONAL CENTERS, SPECIAL ATTENTION MUST BE GIVEN TO THE USE OF PERIPHERAL SERVICES, DEVELOPMENT OF NATIONAL SCIENTIFIC NETWORKS, AND RESEARCH NECESSARY TO IMPROVE SUPERCOMPUTER SOFTWARE PRODUCTIVITY.

WE WILL BE INTERESTED TO LEARN FROM OUR WITNESSES TODAY, THEIR VIEW OF THE ROLE OF THE FEDERAL GOVERNMENT IN ACHIEVING, BOTH SHORT-TERM AND LONG-RANGE LARGE-SCALE SCIENTIFIC COMPUTING NEEDS AND THE CONTRIBUTION NSF CAN, AND SHOULD MAKE, IN THAT REGARD. WE WILL BE PARTICULARILY INTERESTED IN WAYS IN WHICH COMPUTING ACCESS CAN BE ENHANCED, THE ADEQUACY OF BASIC COMPUTER SCIENCE SUPPORT FOR THE DEVELOPMENT OF NEW ARCHITECTURAL ALTERNATIVES AND SOFTWARE, AND THE EFFECTIVENESS OF CROSS-AGENCY TECHNICAL AND POLICY COORDINATION.

Mr. FUQUA. Our first witness is Mr. Henry A. Zanardelli, manager of engineering and product data systems in the computer center at Ford Motor Co. in Dearborn, Michigan.

Mr. Zanardelli, we are glad to have you and delighted to see you.

STATEMENT OF MR. HENRY A. ZANARDELLI, MANAGER OF ENGINEERING AND PRODUCT DATA SYSTEMS, FORD MOTOR CO., DEARBORN, MI

Mr. ZANARDELLI. Thank you very much.

Good morning to the Honorable Congressmen and guests, ladies and gentlemen.

As the chairman has already indicated, my name is Henry Zanardelli. I am employed by the Ford Motor Co. and I manage its engineering computer center in Dearborn, Michigan. From this center, we provided computing services to Ford's worldwide engineering community. Also, we develop computer applications and systems that help our engineers improve their efficiency and the quality of the work they produce. Our center employs a variety of computers, including those in the supercomputer class.

I am pleased to have been invited today to tell you about some of the ways in which we use supercomputers and why they are important to the Ford Motor Co.

Ford is a worldwide leader in the design of automobiles and automotive products. Our mission is to continuously improve our products to meet our customers' needs, allowing us to prosper as a business and to provide a reasonable return for our stockholders.

As you are well aware, the once dominant status of this country's domestic automobile industry has been seriously challenged during the past decade by foreign competition, especially the Japanese. Some of the factors behind this assault are beyond our industry's control, but others are not. Our ability to control two of these factors leading to a loss in dominance—that is, product quality and product cost—is dependent to a great extent on our effectiveness in using supercomputers and related technologies.

The computer is a significant catalyst in helping almost all organizations within Ford do more and better work at less cost, whether the work is technical or administrative in nature. Today, I would like to focus on the use of supercomputers in our product development activities, especially explaining how these computers help us to evaluate the function and the performance of the vehicles we design.

Use of supercomputers for design evaluation usually involves three basic steps.

The first is the development of a mathematical representation or structural model of the part, assembly system, or vehicle being designed. This requires the skills of a specially trained engineer.

The second step involves the processing of the model through computer simulation and analysis programs. Here is where the supercomputer is needed in order to get the job done in a reasonable time.

The third step is the interpretation of the results by an engineer, and at this point the achievement of design objectives is verified, or needed design changes are identified.

The introduction of supercomputers, along with the desire for faster, more cost-effective processing of alternative design studies, has led to the construction of larger and larger mathematical models developed through the use of a technique called Finite Element Analysis, often abbreviated as FEA.

Ford is among the earliest users of FEA. The specific FEA technique we use is based primarily on software originally developed by NASA and called NASA Structural Analysis, more commonly known as NASTRAN. This is an outstanding example of a Government space program byproduct that benefits the private sector.

We use FEA with good results in evaluating the design of individual components as well as large assemblies such as engines, transmissions, suspension systems and complete body structures. The basic concept underlying FEA is that the design of every component or structure can be represented in a model consisting of a series of individual but interrelated finite elements. Typically, a large finite element model will involve from 15,000 to 20,000 discrete elements, with the physical property of each element defined in the model. Then, by means of a supercomputer, trillions of calculations are performed to determine how the design will behave under simulated loads and operating conditions.

A supercomputer capable of hundreds of millions of instructions per second is needed in order to get the job done in a reasonable time. The results of this process are predictions of the deformation, stress, and other physical responses that the component, assembly system or vehicle being designed is likely to undergo in actual on-the-road situations.

In addition to the use of supercomputers and FEA for structural analysis, we apply other computer technology extensively in our efforts to improve vehicle handling, ride, braking, fuel economy, aerodynamic efficiency, emissions, noise and other factors affecting vehicle performance.

The use of supercomputers gives our engineers more analytical insight and time to improve their work. Several design alternatives can be explored before selecting one from which a prototype will be fabricated and tested. And as a result, we reduce the number of prototypes built, we shorten the time to develop new vehicles, we improve product quality, and we save on design and test costs. As one illustration of this, our experience has shown that, when feasible, computer simulation of a design test sometimes can be done at one-fifteenth the cost of a comparable test using an actual prototype.

While FEA and supercomputers already play important roles in Ford's product development process, use of these technologies is still in its infancy. We want to model more complex vehicles and systems. We want to continue maturing in our capability to simulate destructive testing of entire vehicles, expanding on our present capacity for producing both visual and digital output used to analyze crashworthiness of some designs.

We look forward to the day when more automotive styling can be done on computer, and when we can determine the aerodynamic behavior of our car and truck designs without the aid of physical vehicle models and wind tunnels.

There is still a long way to go. Even though we already have experienced some major upgrades, more sophisticated software and even faster supercomputers are needed.

Convinced that the desired tools will continue to emerge in a cost-effective way, we anticipate that our engineering work in the future will involve a paperless process—one through which our designers and engineers will develop and test their designs by computer, storing their results on data bases to be used by the downstream manufacturing activities who fabricate the dies and machinery needed to produce parts for our vehicles.

We already employ long-term training and education programs for our engineers so that they gain an appreciation for computer-based product design and acquire the skills needed to apply it. To some extent, this is a cultural change for them, as they were not exposed to the capabilities of modern supercomputers while in college.

Use of supercomputers in American industry may be restricted, not by hardware or software limitations, but by the limited number of people that have the know-how to use them. Universities and technical colleges need encouragement to adjust their curricula so that there is more instruction in the use of supercomputer technology. More students have to be graduated with an understanding of the supercomputer.

Supercomputers, and the facilities, software, and support personnel they require, are expensive. The acquisition and installation cost easily can exceed $10 million—we know, as we have installed several—and few schools can afford them. Those schools having a supercomputer often tie it up with scientific research projects and only a handful of graduate students gain experience. A means for resource and cost sharing by smaller schools is needed, possibly achieved through the use of supercomputer networks like those provided by some computer service companies.

While the Government is to be congratulated for its recent initiatives to foster supercomputer research at a few large, prestigious schools, I believe more is needed in the way of aid to the smaller institutions who together produce the bulk of our engineering graduates. I encourage you to look into this, as you are today, to identify what needs to be done to assure that our Nation has a continuing and adequate supply of up-to-date engineering talent, and to facilitate the needed programs and incentives.

In summary, I believe the Ford Motor Co. is a leader in the use of supercomputers. It has invested heavily in computers, related facilities, software development and training. I believe Ford will continue to do so in order to produce safe and attractive high-quality products at a competitive cost that customers will want. In my view, the vitality of this Nation, as with Ford, is becoming more and more dependent on the ability of its people to continuously develop and use advanced technical tools such as supercomputers.

Thank you very much.

[The prepared statement of Mr. Zanardelli follows:]

SUPERCOMPUTERS AT FORD

Paper presented to the U.S. House of Representatives
Subcommittee on Energy Development and Applications
and the
Subcommittee on Science, Research and Technology
at a hearing on Federal Supercomputer Programs and Policies
held on June 10, 1985 in Tallahassee, Florida

My name is Henry Zanardelli. I am employed by the Ford Motor Company and I manage its Engineering Computer Center in Dearborn, Michigan. From this Center, we provide computing services to Ford's worldwide engineering community. Also, we develop computer applications and systems that help engineers improve their efficiency and the quality of the work they produce. Our Center employs a variety of computers, including those in the supercomputer class.

I am pleased to have been invited today to tell you about some of the ways in which we use supercomputers and why they are important to the Ford Motor Company.

Ford is a worldwide leader in the design of automobiles and automotive-related products. Our mission is to continuously improve our products to meet our customer's needs, allowing us to prosper as a business and to provide a reasonable return for our stockholders.

As you are well aware, the once dominant status of this country's domestic automobile industry has been seriously challenged during the past decade by foreign competition, especially the Japanese. Some of the factors behind this assault are beyond our industry's control. But, others are not. Our ability to control two of the factors leading to a loss in dominance -- product quality and product cost -- is dependent to a great extent on our effectiveness in using supercomputers and related technologies.

The computer is a significant catalyst in helping almost all organizations within Ford to do more and better work at less cost, whether the work is technical or administrative in nature. Today, I would like to focus on the use of supercomputers in our product development activities, especially explaining how these computers help us to evaluate the function and the performance of the vehicles we design.

Use of supercomputers for design evaluation usually involves three basic steps:

- The first step is the development of a mathematical representation or structural model of the part, assembly, system or vehicle being designed. This requires the skills of a specially trained engineer.
- The second step involves processing of the model through computer simulation and analysis programs. Here is where supercomputers are needed in order to get the job done in reasonable time.

- The third step is the interpretation of the results by an engineer. At this point the achievement of design objectives is verified or needed design changes are identified.

The introduction of supercomputers along with the desire for faster, more cost efficient processing of alternative design studies, has led to the construction of larger and larger mathematical models developed through the use of a technique called Finite Element Analysis, often abbreviated as FEA.

Ford is among the earliest users of FEA. The specific FEA technique we use is based primarily on software originally developed by NASA and called NASA Structural Analysis, more commonly known as NASTRAN. This is an outstanding example of a government space program by-product that benefits the private sector.

We use FEA with good results in evaluating the design of individual components as well as large assemblies such as engines, transmissions, suspension systems, and complete body structures. The basic concept underlying FEA is that the design of every component or structure can be represented in a model consisting of a series of individual but inter-related finite elements. Typically, a large finite element model will involve from fifteen thousand to twenty thousand discrete elements, with the physical property of each element defined in the model. Then, by means of a supercomputer, trillions of calculations are performed to determine how the design will behave under simulated loads and operating conditions.

A supercomputer capable of hundreds of millions of instructions per second is needed in order to get the job done in a reasonable time. The results of this process are predictions of the deformation, stress, and other physical responses that the component, assembly, system or vehicle being designed is likely to undergo in actual on-the-road situations.

In addition to the use of supercomputers and FEA for structural analysis, we apply other computer technology extensively in our efforts to improve vehicle handling, ride, braking, fuel economy, aerodynamic efficiency, emissions, noise and other factors affecting vehicle performance.

The use of supercomputers gives our engineers more analytical insight and time to improve their work. Several design alternatives can be explored before selecting one from which a prototype will be fabricated and tested. As a result, we reduce the number of prototypes built, we shorten the time to develop new vehicles, we improve product quality, and we save on design and test costs. As one illustration of this, our experience has shown that, when feasible, computer simulation of a design test sometimes can be done at one-fifteenth the cost of a comparable test using an actual prototype.

While FEA and supercomputers already play important roles in Ford's product development process, use of these technologies is still in its infancy. We want to model more complex vehicles and systems.

We want to continue maturing in our capability to simulate destructive testing of entire vehicles, expanding on our present capacity for producing both visual and digital output used to analyze the crashworthiness of some designs.

We look forward to the day when more automotive styling can be done on computer, and when we can determine the aerodynamic behavior of car and truck designs without the aid of physical vehicle models and wind tunnels.

There is still a long way to go. Even though we already have experienced some major upgrades, more sophisticated software and even faster supercomputers are needed.

Convinced that the desired tools will continue to emerge in a cost effective way, we anticipate that our engineering work in the future will involve a paperless process -- one through which our designers and engineers will develop and test their designs by computer, storing their results on data bases to be used by the downstream manufacturing activities who fabricate the dies and machinery needed to produce parts for our vehicles.

We already employ a long term training and education program for our engineers, so that they gain an appreciation for computer based product design and acquire the skills needed to apply it. To some degree, this is a cultural change for them, as they were not exposed to the capabilities of modern supercomputers while in college.

Use of supercomputers in American industry may be restricted, not by hardware and software limitations, but by the limited number of people that have the know-how to use them. Universities and technical colleges need encouragement to adjust their curricula so that there is more instruction in the use of supercomputer technology. More students have to be graduated with an understanding of the supercomputer.

Supercomputers (and the facilities, software, and support personnel they require) are expensive. The acquisition and installation cost easily can exceed ten million dollars -- we know, as we've installed several -- and few schools can afford them. Those schools having a supercomputer often tie it up with scientific research projects and only a handful of graduate students gain exposure. A means for resource and cost sharing by smaller schools is needed, possibly achieved through the use of supercomputer networks like those provided by some computer service companies.

While the government is to be congratulated for its recent initiatives to foster supercomputer research at a few large, prestigious schools, I believe more is needed in the way of aid to the smaller institutions who together produce the bulk of our engineering graduates. I encourage you to look into this -- to identify what needs to be done to assure our nation of a continuing and adequate stream of up-to-date engineering talent, and to facilitate the needed programs and incentives.

In summary, I believe the Ford Motor Company is a leader in the industrial use of supercomputers. It has invested heavily in computers, related facilities, software development, and training. I believe Ford will continue to do so in order to produce safe and attractive high quality products, at a competitive cost, that customers will want. In my view, as with Ford, the vitality of this nation as a whole is becoming more and more dependent on the ability of its people to continuously develop and use advanced technical tools such as supercomputers.

DISCUSSION

Mr. FUQUA. Thank you very much for a very interesting presentation.

Since you began, we are very pleased to also welcome Congressman Boucher from the State of Virginia, who is a very important member of our subcommittee.

Mr. BOUCHER. Thank you, Mr. Chairman.

Mr. FUQUA. You indicated on the next to the last page about helping the smaller schools gain access. I assume you are talking about networking and probably royalty fees?

Mr. ZANARDELLI. There are a number of strategies. One is networking, either through a consortium of some sort in the schools, or actually grants of some sort that they can use.

Mr. FUQUA. Do you think there is a role of the Federal Government in trying to provide basic knowledge in the area of supercomputers, particularly software?

Mr. ZANARDELLI. I think there is a lacking in the software and there is an awful lot of work going on in the development of hardware.

The application of that computer to commercial and industrial problems is what is important, and that is one of software development.

Mr. FUQUA. Mr. Lujan.

Mr. LUJAN. Thank you, Mr. Chairman. Just a couple of quick questions.

Also in that same statement, you make note of the fact that we need to insure adequate supply of engineers with the proper knowledge.

Where does a company like Ford, when you get going and expanding, where do you get people to operate these computers?

Mr. ZANARDELLI. Well, we receive most of our scientists and engineers from many local schools. Most are graduates from the University of Michigan. But we have not really hired an awful lot of people who are specialists in supercomputers.

Most graduates have a master's. I think what we would like to see is people coming out in the bachelor's level from the schools who have an understanding of how supercomputers are used.

In many ways you can think of it as private training. Until you get behind the wheel and drive that vehicle, you do not really know.

Mr. LUJAN. You are saying basically you get them from other companies?

Mr. ZANARDELLI. Schools and other companies, yes.

Mr. LUJAN. That leads me to the basic question, I guess, that I wanted to ask. How willing will companies be, companies like Ford—well, specifically Ford—be interested in substantial funding of computer education?

Mr. ZANARDELLI. We are already involved in an internal program of training our people. We do work close with the universities.

We work with Brigham Young, the University of Michigan, Wayne University, in special programs, giving them special jobs that need to be done, having their faculties do them.

Mr. LUJAN. Are those schools near your plants?

Mr. ZANARDELLI. Basically, they are in the Detroit area, but we would certainly be interested in dealing with any school where it could be of use to them.

Mr. LUJAN. Thank you.

Mr. FUQUA. Mr. Sensenbrenner.

Mr. SENSENBRENNER. Thanks.

I would like to follow up on Mr. Lujan's line of questioning.

Let me ask you, sir, how relevant has the education in universities been to people who hire them to do computing for them?

Mr. ZANARDELLI. I will have to say that, in most cases, they have a superficial knowledge. Supercomputing is only one aspect of the whole move toward putting more computer content into the design work. The supercomputer is just one piece of that total puzzle.

It is important for the students to understand how the computer is used throughout the whole engineering process.

Mr. SENSENBRENNER. In other words, what you are saying is that Ford does most of the training of people that they hire, in-house?

Mr. ZANARDELLI. Our designers and engineers ordinarily are trained by us. We have a few specialists that we hired over the years, but the bulk of our designers and engineers are trained in-house.

Mr. SENSENBRENNER. Do you think there are any improvements that could be made to university supercomputing courses that would make the training more relevant and surpass the general and go toward particular needs?

Mr. ZANARDELLI. I cannot give you a specific. I can only give you a generalized answer.

I think, yes, they can be improved. The courses have to provide some hands-on understanding for the typical graduate and they need to make sure that it is not just the elite or the graduate students who have an understanding.

It has to be a commonplace skill, in my view, to strengthen the capability of the American engineering community.

Mr. SENSENBRENNER. Thank you.

Mr. FUQUA. Mr. Boehlert.

Mr. BOEHLERT. No questions.

Mr. FUQUA. Mr. Boucher.

Mr. BOUCHER. Thank you, Mr. Chairman.

I understand that one of the big problems in supercomputing today is software development. And I wonder if you see a significant Federal role in helping to assist software development, and if you do, if you can say a little what that role should be.

Mr. ZANARDELLI. Well, I think that the software that is built, at least in our—I would think the universities could show a strong curriculum and could show a strong facility for use of supercomputers.

Specifically, I cannot identify a specific area. There is such a broad range of involvement. But we do get quite a bit of work involved in the computer design area.

Mr. BOUCHER. Would it be fair to say then that you would foresee a role of the National Science Foundation directly targeted toward supercomputer software development with the work being carried on on university campuses?

Mr. ZANARDELLI. I think that would be a very useful program.

Mr. BOUCHER. Thank you.

Mr. FUQUA. Thank you very much for sharing your views with us this morning. We appreciate it very much.

[Questions and answers for the record follow:]

July 29, 1985

The Hon. Don Fuqua, Chairman
U.S. House of Representatives
Committee on Science and Technology
Suite 2321 Rayburn House Office Bldg.
Washington, D.C. 20515

Dear Mr. Fuqua:

I am pleased to respond to the questions in your letter of June 25, 1985, and appreciate the opportunity to testify before the Subcommittees on Energy Development and Applications and Science, Research and Technology on June 10, 1985.

Question 1. "Will the private sector depend on universities to train supercomputer managers, or would they prefer to hire graduates with advanced degrees in scientific computing and train them to be center managers?"

> Ford would generally prefer the latter approach. Support of the product engineering at Ford requires a number of different information processing disciplines. Our Engineering Computer Center supports data base applications, computer aided design networks, office automation systems, and laboratory automation, in addition to its support of supercomputers. Therefore, our hiring program and much of our internal training are oriented to broad-based technical and business skills.

Question 2. "How did you justify the purchase and operation of a supercomputer to your management? Were you able to quantify real savings of benefits from owning and operating a machine versus leasing supercomputer time from a vendor or using a less-powerful machine?"

> Our supercomputer acquisition was subject to standard Ford procurement practices. It was acquired through a third-party capital lease. Its use over the lease term was projected to improve our product development productivity, primarily through reductions in prototype build, testing and development costs; reduced tooling costs; and reduced vehicle weight and labor costs. The cost reductions were projected to provide a reasonable return on the lease investment. Outsourcing the work to a service bureau was considered, but rejected, due to considerably higher costs than would be incurred internally. Our supercomputer is replacing a less-powerful machine that is no longer able to meet capacity and performance requirements.

I trust that you will find this supplement to my testimony useful to the work of the Subcommittees. Once again, may I say it was a privilege to have participated in the Florida hearing.

Yours very truly,

Henry A. Zanardelli

Mr. ZANARDELLI. Thank you.

PANEL: FEDERAL PROGRAMS AND POLICIES

Mr. FUQUA. Next is a panel of people, and if they will start coming up to the table—Dr. Alvin Trivelpiece, director of the Office of Energy Research of the U.S. Department of Energy; Dr. Mary Good representing the National Science Board of the National Science Foundation; and Dr. Charles Buffalano, who is Deputy Director for Research of the Defense Advanced Research Projects Agency, Department of Defense.

We will hear from all three of you, and then we will have questions.

Dr. Trivelpiece has appeared before this committee many times.

Al, we are very glad to have you, and we will be pleased to hear from you at this time.

[The biographical sketch of Dr. Trivelpiece follows:]

BIOGRAPHICAL SKETCH OF ALVIN W. TRIVELPIECE

Dr. Alvin W. Trivelpiece was nominated by the President on July 8, 1981, and confirmed by the Senate on July 27, 1981, as Director of the Office of Energy Research, U.S. Department of Energy (DOE). In this position, Dr. Trivelpiece serves as technical adviser to the Secretary on the Department's energy research and development programs and is responsible for the multipurpose laboratories and energy education and training activities. In addition, he manages DOE's programs for basic energy research, health and environmental research, high energy and nuclear physics, and magnetic fusion.

Dr. Trivelpiece has extensive experience in the areas of plasma physics and fusion research. From 1978 until assuming his current postion, Dr. Trivelpiece was Corporate Vice President of Science Applications, Inc., of La Jolla, California, primarily responsible for internal research programs relative to innovative technical developments. From 1976 to 1978, he was Vice President for Engineering and Research at Mexwell Laboratories, San Diego California.

During the period 1973 to 1975, Dr. Trivelpiece served with the U.S. Atomic Energy Commission as Assistant Director for Research in the Division of Controlled Thermonuclear Research. Prior to joining the Federal Government, he was Professor of Physics at the University of Maryland, serving in that capacity from 1966 to 1976. From 1959 to 1966, he was a professor at the University of California at Berkeley in the Department of Electrical Engineering.

He received a B.S. degree from California Polytechnic State University in 1953, an M.S. degree in 1955, and a Ph.D. degree in 1958. Both advanced degrees were awarded by the California Institute of Technology.

The recipient of several honors and awards, Dr. Trivelpiece is listed in *American Men and Women of Science, Who's Who in America,* and *Who's Who in the World.* He holds several patents in the area of physics research and is the author and coauthor of over 100 technical reports and books. Dr. Trivelpiece is a member of numerous professional organizations and is a fellow of the American Physical Society, the Institute of Electronics and Electrical Engineers, and the American Association for the Advancement of Science. He has served as a consultant to various private associations and Government agencies and has lectured at conferences and universities throughout the United States and abroad.

Dr. Trivelpiece was born in Stockton, California. He is married and has three children.

STATEMENT OF DR. ALVIN W. TRIVELPIECE, DIRECTOR, OFFICE OF ENERGY RESEARCH, U.S. DEPARTMENT OF ENERGY, WASHINGTON, DC

Dr. TRIVELPIECE. Thank you, Mr. Chairman. It is a pleasure to be here.

I would like for my prepared statement to be made part of the record, with your permission.

Mr. FUQUA. Without objection, it will be. If you want to just summarize, that will be fine.

Dr. TRIVELPIECE. One of the reasons it is a pleasure to be here is that field hearings usually are a result of something that has come up in the way of the press calling attention to environmental problems the Department of Energy has caused, or there is a laboratory which happens to be currently underfunded in some outrageous manner. So something is usually wrong.

This time, I am pleased that something is right to cause the field hearing. So it is with some sense of pleasure that I am here today.

I have testified before this committee on this subject about 2 years ago and I am pleased to have this opportunity to update things.

I would like to begin with a brief bit of history here. It was in 1973 I was working for the Atomic Energy Commission, and at that time commissioned a study on what its needs were for scientific computing. Bennett Miller, who worked for me at that time, led that study.

At that time, the DOE Fusion Program was using the equivalent of one 6600 in the collection of problems—that is the entire Fusion Program throughout the country was using the equivalent of one 6600.

The work that was done indicated that the collection of problems that needed to be addressed and the effort that would be required to solve them was going to require something well in excess of a 7600.

Well, it you look at the problems then and the way things are now, it is kind of modest when you realize what is happening. At that time, when we realized this, the thought was that we would like to make the 7600 available. We had four principal laboratories and the idea of making four computers available was not going to be supportable.

At that time, we came up with the idea that we would try to locate one central computer and connect it through a communications network to the other laboratories. I was assured by a large block of experts that that was absolutely impossible to do in a practical way.

Fortunately, not being that much of an expert myself, I ignored their advice. But there was very little support in the Atomic Energy Commission for this idea.

Fortunately, there was support by at least one person and that was the controller of the AEC, who had a substantial amount of influence on these things. The reason that it worked was I think in part due to his support from the beginning.

The things that have happened since then, however, are the result of a lot of hard work on the part of many people who have done a very good job. Some of those people are here today and one I would like to mention is John Killeen, who certainly played a key role, and many others that I will not mention, in establishing the fusion computer center to establish a model that seems to work.

I notice now that others are beginning to use this model as a means of establishing computer centers and networks that support them.

I would like to turn now briefly to national security. National security takes many forms and is used in many different ways.

Secretary Herrington regards energy and energy technology as really essential elements in the national security enterprise in the United States, not in the narrow sense that it is a military activity, strictly speaking, but in a broader sense that science and technology and economy are all essential elements in our competitive structure. And we maintain our edge in every respect.

Part of that edge, our national competitive edge, involves the use of advanced computer systems, whether it is in looking for oil reserves or looking at fusion research or economic modeling or such things recently as the Grumman Aerospace Corp.'s design of the X–29 aircraft, which is a remarkable device, and the design of nuclear weapons.

Well, in that sense, DOE has played, and will continue to play, an important role in the development of a supercomputer and this is by virtue of seeking to meet its own needs to design weapons and to do advanced scientific programming.

DOE at the moment has something like 22 advanced computer systems, and I believe is the largest user.

When I joined the Department of Energy for this tour of duty with the Government, I commissioned another study, which, interestingly enough, was also done by Bennett Miller. This study indicated that there were other areas within the scientific enterprise department that did require supercomputer capability, and on a somewhat urgent basis, in fact.

Well, the basis was sufficiently urgent that I did not believe that we could go through the procurement cycle with sufficient speed to make this access available, so I allocated 5-percent time on the Magnetic Fusion Computer Center and made it available to the other areas for nonmagnetic fusion users within the Department.

This utilization was expanded with the siting of the Cray XMP–22 at the Fusion Computer Center and the 1985 allocation of that time has something on the order of 5,000 hours for high energy physics and about 5,000 hours for basic energy sciences and 1,000 hours for the biological sciences which the Department supports.

It is worth noting that within the first month that the computer was up and running, 90 percent of its capacity was utilized and, in doing so, stipulated the problems that be addressed with this device not simply be the mundane routine type of computations that could be done with minicomputers but rather for those problems for which progress cannot be made without adequate access to the classic systems and beyond.

This activity is now being expanded. We have new nodes at the Stanford Linear Accelerator Center, at Brookhaven National Laboratory, at Fermi, the Oak Ridge National Laboratory and at Ames as well, and recently, Mr. Chairman, has been connected to the Florida State University Computer Center here and the Cyber 205. I think some of those technical details are going to be covered by Dr. Killeen.

The access is, even for the Department of Energy, is not yet what is needed, but it is substantially better and we are making progress. Unfortunately—or fortunately, depending on how you

look at these things—DOE is really not in a position to provide access to the general community of unversity users.

There are some aspects of the Brooks Act and other dealings with Government regulations that make it impossible for us to be able to do that in the same way the National Science Foundation can, and we are pleased that the National Science Foundation is moving in to provide the kind of access to the academic community that the Department is unable to do.

Just let me turn briefly from the access problem to research on computational processes. We have tried to expand that activity within the Department and we have made some progress. There are a number of details in my prepared testimony on the activities that the Department is supporting, parallel processing of research and so on.

We plan to continue to expand the access to supercomputers to DOE scientists and engineers in several areas, and biology is one which I hope that we most rapidly grow in.

I find it fascinating that biological science has been somewhat behind the power curve, both in supercomputer and ordinary computational sciences and believe that they can benefit greatly from that. Not just simply gene banks and things like that, but the modeling of very complex biological molecules.

Supercomputers are an indispensible part of the U.S. science enterprise, and we need to continue to expand our use in many fields and try to provide training in their use. The ability of the United States to compete in the world, I believe will depend in part on how well we end up doing this.

Finally, Secretary Herrington, in responding to the President's desire to help in the high school education arena, asked me what we might do in addition to what we were already doing in our laboratories in terms of providing an opportunity for high school students to spend part of the summer or other periods of the year in the lab.

Several suggestions were made and some of these are actually being implemented. One in particular that I want to call attention to was that Secretary Herrington sent a letter to the 50 Governors of the States as well as Puerto Rico and the District of Columbia, and in that letter he cited the President's ongoing commitment to education and high school students and so on.

But the thing he did was also invite each one of those States and the District and Puerto Rico to send one student, or to nominate one student, the best student in the State, in mathematics, physics, science, computational science, and we would take them at the Lawrence Livermore National Laboratory for a 2-week period in the summer, which ends on August 14, I believe. They will spend 2 weeks out there getting their hot little hands on a supercomputer, probably for the first time.

We look forward to this, and it has a great deal of support. We have all students nominated now. I am sorry I did not bring with me the name of the individual who has been selected from the State of Florida, but I am sure that can be obtained if you need it.

The thing that will have to happen, however, when these students have left the MFC at Livermore, is that all the passwords will have to be changed.

Thank you, Mr. Chairman.

[The prepared statement of Dr. Trivelpiece follows:]

Statement of Alvin W. Trivelpiece, Director of Energy Research, Department of Energy

Mr. Chairman and Members of the Committee:

I am pleased to have the opportunity to testify before you on supercomputers. Supercomputers are generally defined as the most powerful computers available at any given time for solving large scientific and engineering problems. These machines are capable of performing hundreds of millions to over a billion arithmetic operations per second and have several to hundreds of millions of words of internal memory.

I testified before this Committee on the subject of supercomputers almost exactly 2 years ago. Today I will bring you up-to-date on the Office of Energy Research's (ER) program in the areas of access to supercomputers and long-range computational researcn.

In recent years, the modern supercomputer has changed the nature of scientific research and technology development. The advancement of science used to depend upon experiments for data and on theory to gain understanding. Today there is a third equally important ingredient to scientific research, computational science. Computational science serves a role that is a hybrid between that of theory and experiment. In some cases computations provide insights into experimental data, and in others, computations are used to simulate the ideal experiment to test an analytical model. The emergence of computational science as a important element in scientific research and technology development is the result both of the development of our ability to do computational modeling of complex physical problems and of the enormous power of the modern supercomputer. This combination allows scientists and engineers to model complex problems in a much more realistic way and to obtain much more accurate answers than was possible just 5 years ago.

The capabilities of present day supercomputers to accurately model very large problems in scientific research and engineering development have made the supercomputer an important scientific research tool both for the Federal government and industry. In fact, it is often said that today's supercomputers are rapidly becoming the "machine tools" of the modern high technology world. There is little doubt that if the U.S. is to maintain its leadership in research and development, there will need to be a continuing supply of more advanced supercomputers available to our research community and individuals trained to both run and effectively use them. There are already many government research and development programs that are highly dependent on supercomputers. These programs include nuclear weapons development, cryptology, fusion research, aeronautics, and weather forecasting. In addition, supercomputers are being used more and more for engineering in a number of industries such as petroleum, electronics, aerospace, automotive, chemical and even movies. In addition, this computer capability is increasingly useful in many areas of basic scientific research.

Within the U.S., the Department of Energy (DOE) is the largest user of supercomputers with a total of 22 systems currently installed. Many of the Department's technical programs face difficult scientific and engineering challenges in a variety of areas, such as the design of nuclear weapons, the behavior of hot plasmas in fusion research, nuclear reactor safety, carbon dioxide research, materials science, combustion research, and nuclear physics. DOE's involvement with supercomputers is not new. The Department's laboratories, Los Alamos National Laboratory (LANL) and Lawrence Livermore National Laboratory (LLNL), historically have been on the forefront of

supercomputer development because of mission requirements. These laboratories have obtained the first, or one of the first, of each new generation of supercomputers. In each case, the computer has been delivered almost devoid of software, and the laboratory was required to develop most of the necessary software for a complete system. Thus, the Department has been a key element in the development of this technology in the past and continues in that role today.

Earlier, I indicated that there have been great advances in computational science and that modern supercomputers are powerful enough to "solve" many complex problems. While it is true that great progress has been made, there is still much more to accomplish. Many of the Department of Energy's and, indeed, the Nation's research and development programs have large, complex scientific problems to solve that can be only crudely approximated with today's supercomputers and mathematical techniques. There is a need to include more "physics" in the computational models, to use finer zoning for more accurate answers, and to solve full three-dimensional problems. In order to solve these complex problems to the accuracy desired in the future, computers with substantially increased computing power over those currently available will be required, computers that are hundreds to thousands of times more powerful than today's machines. Limitations on useable computing power are difficult to see at this time.

Applied Mathematical Sciences Research Program

Because of the importance of supercomputing to accomplishing the missions of many of the Department's programs, the Department established a new organization to support several aspects of computational science. This new

organization, the Scientific Computing Staff, reports directly to me. The objectives of this organization are to meet the immediate needs for supercomputer access by the research programs supported by the Department's Office of Energy Research and to meet the long-range needs of the Department in computational research.

Energy Science Advanced Computation

The goals of this program are to provide access to modern supercomputer systems for researchers that are funded through the Office of Energy Research programs and to prepare this research community to take full advantage of new supercomputer technologies as they become available.

In May 1983, I responded to several studies and reports that cited the need for expanded access to supercomputer resources for the U.S. scientific research community by initiating this supercomputer access program. To get this program moving quickly, I allocated 5 percent of the resources at the National Magnetic Fusion Energy Computer Center (NMFECC) to scientists in the non-Magnetic Fusion Energy (MFE) ER programs who had previously had limited or no access to supercomputer systems. This allocation was less than 10 percent of the total hours requested by this community of users. Nonetheless, feedback from these new users was very favorable and the resource allocations were effectively used. Because of the success of this program, again in FY 1984 the same amount of computer time was made available to this community.

During the first quarter of FY 1985, the advanced computation program moved into full operation with the installation of a Cray X-MP22 computer system at the MFE Computer Center. This system successfully passed acceptance testing in December 1984 and has performed extremely well to date. Representatives from all the ER program areas and the Scientific Computing Staff established resource allocations based on programmatic needs and scientific merit. These allocations have been implemented on the Cray X-MP22 computer system by the NMFECC staff and are summarized as follows:

ER Program	FY85 Hours Requested	FY85 Hours Allocated
High Energy and Nuclear Physics	10,090	5,350
Basic Energy Sciences	9,020	4,480
Applied Math Sciences	3,478	1,485
Biological and Environmental Research	1,256	700

Both the number and size of the proposals requesting supercomputer time verify the substantial need for supercomputer resources in all aspects of energy research. Implementing these allocations has added over 1200 new users at the NMFECC. (Approximately 37 percent of the resource allocations were made to university research projects.) During December 1984, the first full month of installation, this Cray X-MP22 computer system was over 90 percent utilized. This high utilization level was possible as a result of the several studies that accurately defined our needs and our use of an existing, properly staffed supercomputer center with a nationwide communication network. Currently the machine is operating at full capacity.

Later this month, we will complete the installation of the new MFE network nodes at the DOE laboratories that have been awarded large amounts of computer time. These nodes located at the Stanford Linear Accelerator Center, Lawrence Berkeley Laboratory, Fermi National Accelerator Laboratory, Argonne National Laboratory, Brookhaven National Laboratory, Oak Ridge National Laboratory, and Ames Laboratory will further facilitate use of the Cray X-MP22 system by the non-MFE user community.

In October 1984, ER entered into a cooperative agreement with Florida State University (FSU) to establish a Supercomputer Computational Research Institute (SCRI). The SCRI will conduct research in computational science related to the ER mission, will develop advanced software tools and techniques for use on supercomputer systems, and will operate a supercomputer facility to provide services to FSU, to the SCRI, and to some ER users. The SCRI installed a CDC Cyber 205 Computer system. This system was recently incorporated into the nationwide MFE network. Allocations have been established and implemented for these resources in the same manner that resources were allocated at the NMFECC. Over 80 percent of these resources have been allocated to university researchers.

In 1983, the Office of Science and Technology Policy (OSTP) formed a Federal Coordinating Council on Science, Engineering, and Technology (FCCSET) Committee to examine what the U.S. should do to maintain its leadership in the development and use of supercomputer technology. This Committee formed three subcommittees to develop recommendations on the role of the government in maintaining the U.S.'s lead in supercomputers, on providing access to

supercomputers to more researchers, particularly in universities, and on coordinating the roles of each agency in the area of research on high performance computing. Each of these subcommittees produced a report and continues to exist in a modified form to provide coordination between agencies and to provide policy recommendations to OSTP.

The Department of Energy is participating in the FCCSET subcommittee on Supercomputer Data Communication Networking along with the Defense Advanced Research Projects Agency, the National Science Foundation, the National Aeronautics Space Administration, and the National Bureau of Standards to coordinate individual agency data communications plans and activities. Coordination and cooperation in this area emphasize the need to document individual network development projects, to position agencies to promote and to take advantage of the emerging data communications standards, and to interface with external data communications networks. This cooperation is especially important to facilitate interagency and/or inter-university research collaborations that are dependent on supercomputer support.

Computational Science Research

The goal of the Computational Science program is to advance our understanding of the fundamental concepts of mathematical models of the key physical processes under investigation in the Department's research and development programs. Much of the scientific research and development effort throughout the DOE programs is focused directly on analytical and numerical modeling of physical processes. Basic research in the mathematical formulation of scientific and engineering problems and the creation of algorithms for solving

these problems on novel computer architectures with many parallel processors will make a significant contribution to the development of future energy production and distribution systems.

The move to parallel processing is a dramatic departure from the sequential and vector processing of the last thirty years and is necessary because there are limits on the extent to which current processing components can be improved. To achieve the needed increases in speed, more attention must be paid to parallelism and problem decomposition. The efficient use of these parallel machines is intrinsically more difficult than sequentially structured machines. This program emphasizes research on concurrent algorithms and software to exploit these new architectural ideas. Availability of these methods will set the pace for application of the next generation of supercomputers and will probably influence the ways in which they are designed to operate.

This program supports interdisciplinary teams including analysts, applied mathematicians, computational scientists, postdoctoral fellows and graduate students collaborating in universities and national laboratories to carry out research on parallel algorithms for mathematical software tools. As documented in the David Committee's report, Renewing U. S. Mathematics: Critical Resource for the Future, published by the National Research Council of the National Academy of Sciences, "DOE's research program in the applied mathematical sciences has been and continues to be the leading federal agency program in support of research at the interface between the mathematical and the computational sciences." The intent is to strengthen the ties between universities and the Labs and to provide salary support and

computational resources in the disciplines associated with computational science. The DOE program, with its unique history of basic research in these areas, is positioned to provide strong leadership in filling the "pipeline" from graduate studies to postgraduate studies to research positions in academia and the laboratories.

Recently, we surveyed a couple of dozen promising ideas on parallel architectures that are scalable to massively parallel supercomputers. These fall into two broad categories: shared memory, where many processors share one large memory; and distributed processors, each with a local memory and some form of communication among them.

The shared memory concept is implemented in the latest supercomputers available from U.S. computer vendors and is key in several research ventures having partial DOE support, such as the Ultracomputer project (a joint effort with New York University's Courant Institute and IBM Yorktown Heights) and the CEDAR project (at the University of Illinois).

Distributed processor systems range from collections of scientific workstations on local area networks (such as DOE funded efforts at the University of Florida, Lawrence Berkeley Laboratory, and the University of Wisconsin) to machine designs like large-scale hypercube architectures (pioneered by the group at Cal Tech with partial support from DOE and now available from at least three computer companies).

Research projects supported by the program address the issue of providing all levels of software for novel parallel architecture machines, from the basic operating system to structuring compilers for high-level languages, to developing sophisticated mathematical algorithms for using these machines effectively. Other projects supported by this program are using commercially available parallel machines to begin to address these issues. DOE groups at Argonne, Oak Ridge, Los Alamos, Lawrence Livermore, and several university contractors have already reported promising results in a variety of algorithms tailored for parallel implementation in existing scientific computer programs. Research projects started this year include as a necessary ingredient interdisciplinary teams of scientists addressing the full range of issues required to solve computational problems, from analytic methods to high-level languages, to algorithms for parallel architectures. Examples of these projects include the Supercomputer Computational Research Institute at Florida State University, the collaboration between Argonne National Laboratory and the University of Illinois, and the Cal Tech group.

The Department of Energy is the lead agency in research and applications of large-scale scientific computation. The DOE program has not concentrated on issues related to special purpose machines such as those adapted for signal and image processing or various artificial intelligence applications. These areas are the major focus of the Department of Defense's strategic computing and basic research programs. We monitor progress of those and other agencies' research programs by participating in the FCCSET subcommittee on High Performance Computing and the Interagency Committee on Extramural Mathematics Programs (ICEMAP). DOE is also one of the ICEMAP sponsors of the newly formed Board on Mathematical Sciences in the National Research Council of the National Academy of Sciences.

Cooperation with U. S. industry is an important element in many federally-funded research efforts. The FCCSET subcommittee on supercomputer access maintains contact with industry representatives. Technology transfer of research results to industry has been effective. The commercialization of the Cal Tech hypercube and of the University of Colorado/Argonne Toolpack software development environment package as well as the adaptation of quality mathematical software packages for the latest supercomputers are examples of this interaction. Industry participates with our researchers by donating components, providing computer time on prototype supercomputers, and offering substantial discounts to universities and laboratories on newly developed parallel machines.

Future Plans

The Department's plans for the future in the supercomputer area consist of three parts: providing high quality supercomputer service to those DOE supported researchers and engineers who require such computing resources; keeping our supercomputer centers supplied with the most powerful hardware available; and supporting the computer science and applied mathematics research required to make efficient use of both available and next generation supercomputers.

The Department believes supercomputers are an indispensable scientific and engineering tool required to help meet DOE's mission objectives. Continued access and research will also contribute to maintaining the Nation's leadership in this field of supercomputer technology. We hope to maintain momentum through the difficult budget constraints facing us in the next few years and continue to improve the vitality of our computational capabilities.

Mr. Chairman, this concludes my prepared statement. I would be pleased to answer any questions at this time.

Mr. FUQUA. Thank you very much, Dr. Trivelpiece.
Dr. Good.

STATEMENT OF DR. MARY L. GOOD, CHAIRPERSON, PROGRAMS AND PLANS COMMITTEE, NATIONAL SCIENCE BOARD, WASHINGTON, DC; AND PRESIDENT AND DIRECTOR OF RESEARCH, SIGNAL UOP RESEARCH CENTER, DES PLAINES, IL

Dr. GOOD. Thank you very much, Mr. Chairman. I am very pleased to be here this morning.

I have a prepared text, as you know, and I would simply like to ask that that be made a part of the record.

Mr. FUQUA. Without objection, it will be.

Dr. GOOD. I would simply like to paraphrase part of that.

As the text states, I actually am president and director of research at the Signal Research Center in Des Plaines, so I have a similar background as the gentleman from Ford this morning. I really appreciated his comment.

However, today I am here in a different role and that is as chairwoman of the Programs and Plans Committee of the National Science Board, where most of the discussion of the supercomputer program takes place.

I am very pleased to be here because I think it is important that we give a brief overview of what the initiative at the NSF has been and where we are today, and talk about, at least in some detail, the places that we believe we are going to be able to make an impact.

Now, the initiative, as you know, at the Foundation is relatively new and the objectives of that initiative are really threefold.

The first one, as far as NSF is concerned, is the possibility of really moving the science forward. In other words, the ability to give our research people the opportunity to solve some scientific problems which they have not been able to do before, which obviously has to do with the science phase.

The second one, which has been very much discussed earlier, is the whole concept of training students, and so the second major goal is to provide the opportunity to train a new generation of students and also beginning research people in the use of advanced computational instrumentation.

The last one is that we hope that the NSF initiative, and we are trying to leverage this particular position, we hope that that initiative will stimulate the computer industry itself, so that we will in this country develop better the next generation of supercomputers, and perhaps this utility will push that in that direction.

Now, if we look in detail about the way that program has been put together, the first discussion really within the NSF started about 5 years ago.

And the comments that were made earlier from one of the Congressman about Dr. Wilson's inability to get time on supercomputers in the United States, that discussion poured over into the NSF, as Ken very well knows, and that was just about 5 years ago at this moment.

Now, the principal findings of all of those discussions was that the academic community did indeed have very little access to advanced computing equipment. In other words, the supercomputer,

one of the more major tools today, was not generally available for advancing the frontiers of science and perhaps even more important, or at least equally important, there were very few, if any, students being trained with hands-on access to this equipment.

So the recommendation was that the Federal Government establish a network of supercomputers which would provide that kind of access to researchers in a wide variety of fields.

Now, just to give you where we are, the response from the NSF to these recommendations in late 1983, NSF reserved $6 million out of its fiscal year 1984 budget to start that initiative. This was money that was actually moved from other programs. A few months later, a request for $20 million was included in the fiscal year 1985 budget to set up an initial center and network. In April of 1984, the Office of Advanced Scientific Computing was established with Dr. John Connolly as the director for that.

He is here today. So if we get into any technical components, we can ask him to respond.

The congressional action, partly as a result of your committee, Mr. Chairman, modified that 1985 budget to $40 million, and so the plan was modified to include startup funds for three supercomputer centers.

Now, the first contracts were awarded in August of 1984. These went to the Boeing Computer Services, the University of Minnesota, and Purdue. Those contracts purchased the equivalent of about 5,200 hours of supercomputer time, which was distributed to academic research people on the basis of merit, via peer review requests of proposals that came in.

Now, since that time, over 300 research groups from more than 80 institutions have been connected with one or more of those three centers. So the response time, if you consider August 1984 is not that far back, the response has been really just extraordinary. I cannot think of any program NSF has had where the response has been as rapid and where the—the slope of the curve is unusual.

The next three, the time purchased at existing centers, will now be expanded. We have three other centers that have been recently started. One is at Digital Productions, the Colorado State University system, and the AT&T Bell Systems Laboratories.

Then, in addition, the congressional group last year became aware of the Cray supercomputer that is at NASA Lewis and asked us to find an appropriate home for that. That is underway at the moment, and I suspect that a final decision on that will be made at the June National Science Board meeting. As I understand it, the recommendations are in order to do that.

Now let me say a few words about the networking side of the initiative, because I think this is really one of the important issues. At the moment, it is functioning very well, but the Foundation really feels that it is still considerably less than adequate. For example, if a researcher is not fortunate enough to be connected to a major network, he ends up with a very slow information transfer system, which is really not very suitable for advanced scientific computing. Now, in the short term, we think we can improve that situation by enhancing and alternating existing networks, and that is going to be the first step. Now, some of the options are currently being considered: one is expanding the ARPANET nodes. Again,

that particular consideration is going to come up at the June Board meeting. We believe we are going to be able to move with that fairly quickly. The other is trying to improve the campuswide access to existing nodes. And, last, in the longer term, is to attempt to upgrade other networks, particularly things like BITNET and CSNET and installing dedicated links between the users and the centers.

Now, the BITNET line is very important. There are about 200 users on that already and that is the one in many ways which is accessible to the type of institutions that we discussed earlier. The smaller universities can afford that and can get on the system through that network, and the possibility of expanding that is where we would like to go in addition to the ARPANET as well.

The Foundation has also attempted to coordinate as well as it can with other agencies, and we are in the process of trying to put together some workshops. Three of those—one at the University of Minnesota, one at Boeing Computer Service, and one at the National Center for Atmospheric Research in Boulder—will be held actually this summer. They are jointly sponsored by the National Science Foundation and the Department of Defense. Those, I think, are going to have a major impact again on users. I think those are the kinds of things that are going to build a user community.

Now, the immediate plans of the NSF include three new supercomputer centers. These will be dedicated primarily to NSF-supported research. Those have now already been announced and contract negotiations are underway with the three units, and those should be up and running very shortly. One is the San Diego Supercomputer Center. The second is the John von Neumann Center at Princeton. And a third one is at the University of Illinois, the one which already has a Cray in-house, an XMP, and we hope to upgrade that.

Now, with respect to the immediate plans on the networking side, we do hope to have an NSF nationwide network which will provide megabaud access for about 100 institutions to approximately 10 supercomputer centers. We would expect and hope to have that up and running within about 3 years.

The other areas that we are looking at, besides the centers and the networking programs, the others we are in the process of supporting through other disciplinary programs at the Foundation. The first of these is experimental computer systems. This was discussed earlier—the possibility of generating the next generation of supercomputers here with people who have state-of-the-art information they would like to develop in that area.

The second one is perhaps one that has been discussed and is really an extraordinarily important one, and that is software development and computational mathematics. We are attempting to get larger programs in those areas through the various disciplinary pieces of the Foundation.

And then, of course, the last one is local facilities for access to the networks. One of the problems of getting people on the network is that many local universities need to be able to provide proposals just to buy the front end so that they can get on the network. This is a small grant but very important because many universities have to have that to be able to participate.

So, in conclusion, I would like to point out that it really is kind of a critical time for the computational science community because most academic scientists, even with the improvements that I have just talked about, most academic scientists are still tailoring their problems to a class 4 or 5 computer when class 6 machines have already become obsolete. The machines available now are class 6.5 and the next generation of class 7's are expected to be on the market next year.

So how we keep up with this rapidly moving front is a little difficult, but we are going to have to try.

We think that the NSF program represents an important step in providing not just access for a few researchers, but in providing a research environment in which computation will advance science and engineering across the entire spectrum of disciplines.

I thank you very much, Mr. Chairman, for the opportunity to participate, and I will answer any questions later.

[The prepared statement of Dr. Good follows:]

Statement of Dr. Mary L. Good, Chairperson, Programs and Plans Committee, National Science Board

TESTIMONY ON
THE NSF SUPERCOMPUTER INITIATIVE
BEFORE THE HOUSE SCIENCE AND TECHNOLOGY COMMITTEE

DR. MARY L. GOOD
CHAIRPERSON, PROGRAMS AND PLANS COMMITTEE
NATIONAL SCIENCE BOARD
JUNE 10,1985

Good morning. My name is Mary Good and I am President and Director of Research at the Signal UOP Research Center in Des Plaines, Illinois. However, today I am testifying as the Chairperson of the Programs and Plans Committee of the National Science Board.

It is a pleasure for me to appear before this committee today to give you a report on the NSF supercomputer initiative which, I think most people would agree, has made a considerable amount of progress over the past year.

This initiative has received a great deal of attention both inside and outside NSF, as it is seen as a major step in meeting several national needs:

First, as a means to make a significant impact on the nature of scientific and engineering research in this country, we expect this initiative to enable our researchers to solve many important unsolved problems.

Second, this program will train a new generation of students and beginning researchers in the use of advanced computational instrumentation. These young scientists and engineers will be needed to use the supercomputers which are already vital to the research and development efforts in industrial and government laboratories.

Third, we anticipate that the NSF initiative will have an important stimulative effect on those innovative parts of the computer industry which are likely to develop even better supercomputers in the future.

These are all major factors which are needed to maintain our leadership role in technology, and to meet new challenges.

A general theme of the NSF supercomputer program is a recognition that computation has recently become the third branch of scientific endeavor which has emerged to complement the other two branches, namely theory and experiment. Indeed, in many areas of science and engineering, one sees computer simulation as a

substitute for some experimentation. It has become possible to forego many aspects of experimental testing, by replacing them with the faster (and often cheaper) supercomputer simulations. The result is that supercomputers often lead to a significant increase in scientific and technological productivity.

In the rest of this presentation, I will outline the recent history of the NSF initiative, describe its current status, and discuss its plans for the future:

Almost five years ago, discussions began within NSF on the need of scientists and engineers for obtaining access to the best possible computational equipment.

The Press report (1981) to the NSF Physics Advisory Committee was an early result of these discussions. It recommended that NSF set up a network to connect to a supercomputer center dedicated to computational physics. At the time, its recommendations were regarded as premature and too costly for an initiative in only one discipline.

The subsequent study known as the Lax report (1982) explored the computational needs of all branches of science and engineering.

The principal finding of these various studies was that the academic research community had very little <u>access</u> to the most advanced computing equipment (i.e. supercomputers and associated peripherals) except for a few specialized fields (e.g. atmospheric science, plasma physics and aerodynamic simulation).

In other words, an important tool, the supercomputer, was not generally available for advancing the frontiers of science across a wide variety of disciplines. In addition, those reports noted, the academic community was losing its computational talent to the industrial and governmental laboratories. This internal "brain drain" restricted the ability of our educational system in preparing the next generation of scientists and engineers to meet future challenges.

As a result, the Lax report recommended that the Federal government should establish a network of supercomputers which would provide access to researchers in a wide variety of fields.

The response of the National Science Foundation to these recommendations started in late 1983 when NSF reserved $6 million out of the FY84 budget to start the initiative. A few months, later a request for $20 million was included in the FY85 budget to set up an initial center and network. In April of 1984, the NSF established the Office of Advanced Scientific Computing, reporting to the Director, under the direction of Dr. John Connolly. In June of 1984, Congressional action increased the FY1985 budget to $40 million. The NSF plan was then modified to include start-up funds for three supercomputer centers.

In August of 1984, NSF awarded three contracts for services at existing supercomputer centers. These were with Boeing Computer Services, the University of Minnesota and Purdue University. These three contracts purchased the equivalent of 5200 hours of supercomputer time, which was to be distributed to academic research projects on the basis of merit. This was felt to be the best way of meeting the need for access until the new NSF supercomputer facilities could be established.

Since that time, over 300 research groups from more than 80 institutions have been connected with one or more of the three centers. Research groups in physics, chemistry, materials research and engineering have been dominant. Other fields, like biology and geology have expressed interest in using supercomputers, but have been slower in getting started up. However, the potential for problem solving in these areas is just as exciting as in the more computationally-oriented disciplines. The time available to the community will be expanded in a number of ways to meet the anticipated demand:

First, the time purchased at the existing centers will be substantially increased. Second, three other centers have been recently started,- at Digital Productions, Colorado State University and AT&T Bell Labs. Third, as mandated by Congressional action last year, the Cray supercomputer, currently being used at the NASA Lewis Research Center, will be transferred to an institution of higher education. This institution will be chosen by NSF on the basis of merit.

These interim centers will fill the projected demand for high-quality science and engineering projects for the next year or so when it is expected that the NSF supercomputer facilities will be ready to meet the needs of the community.

As to the networking side of the NSF initiative,- while the centers are functioning well, we still consider the remote access to be less than adequate. Typically a researcher not fortunate enough to be connected to a major network has to be satisfied with an information transfer rate of only 1200 bits/second. This rate is satisfactory for limited textual messages, but is impossibly awkward for the transfer of average computer files. In the short term, this situation will be improved by enhancing and augmenting existing networks. Some of the options currently being considered include: expanding the number of ARPAnet nodes; improving campus-wide access to existing nodes; upgrading other networks such as BITnet and CSnet and installing dedicated links between users and centers.

At the minimum level, we expect to fund our existing centers to link to those networks which are useful to the scientific community. We hope that, within the year, we will have upgraded the quality of access for most researchers by at least an order of magnitude.

We note that all of the NSF activities are being discussed with those of other federal agencies through FCCSET (Federal Coordinating Committee for Science, Engineering and Technology). At several meetings a year, supercomputer and networking plans are exchanged and coordinated.

In addition, in coordination with other agencies, a series of supercomputer institutes will be held this summer. Three of these, at the University of Minnesota in Minneapolis, at Boeing Computer Services in Seattle, and at the National Center for Atmospheric Research in Boulder, Colorado will be jointly sponsored by NSF and the Department of Defense.

These summer supercomputer institutes will be designed to introduce students and new researchers to the latest supercomputer equipment, as well as serving as a forum for exploring new applications and computational developments.

All of the preceding discussion has to do with what has been accomplished by the NSF programs so far this year.

The immediate plans of the NSF initiative include the establishment of three new supercomputer centers, which will be dedicated to NSF-supported research.

As a result of a nationwide competition, and an extensive peer review process, three institutions have been chosen to be the sites of these centers. These centers have been approved by the National Science Board, and the awards will soon be made.

The following centers will be funded:

(1) The San Diego Supercomputer Center, which will be on the La Jolla campus, and will have a Cray XMP configuration. This center will be managed by GA Technologies;

(2) The John von Neumann Center which will be near Princeton, NJ and will have a CDC machine, upgradeable to an ETA-10; and

(3) U. of Illinois, which will start with a Cray XMP and upgrade every year in a modular fashion.

Current plans anticipate the establishment of more centers in the next two years, if budgets allow. We expect the NSF portion of the support of each of these centers to be about $10 million/year in 1984 dollars.

On the networking side, a planning process is now being completed to establish an NSF Nationwide Network, which will provide megabaud access for approximately 100 institutions to approximately 10 supercomputer centers. We anticipate that this will be accomplished within three years.

Experiments will be started this year to test the technology necessary for the National Network, for example, satellite transponders, earth stations, and fiber optics.

Besides the centers and networking programs, NSF will be supporting other areas, including research designed to take advantage of new technologies as they develop. For example:

(1) Experimental computer systems: These are machines which are not yet on the commercial market, but are useful for research purposes. The National Science Board at its February meeting approved the "hybrid supercomputer" proposed by Cornell, which will consist of a large IBM mainframe connected in parallel to several FPS array processors.

(2) Software development and computational mathematics. Support for these areas of research is designed to tap the talent in computer science and mathematics for problem optimization. Experience has shown that a large percentage (perhaps half) of computational productivity advances are attributable to improvements in numerical algorithms and fine-tuning computer codes. This will be particularly important for vector and parallel machines.

(3) Local facilities for access to the networks. This will include local workstations, input-output devices, graphics facilities, etc. Eventually local facilities could include distributed supercomputers as they become available, while the centers would be reserved for the next generation of super computers.

In conclusion, I want to point out to you that this is a particularly critical time for the computational science community. At the present time, most academic scientists are tailoring their problems to a Class 4 or 5 computer when Class 6 machines have already become obsolete. The machines available now are class 6.5, and the next generation of class 7's are expected to enter the market next year.

We believe that the NSF program represents an important step in providing not just access to supercomputers for a few researchers,.. but in providing a research environment in which computation will advance science and engineering across the entire spectrum of disciplines.

Thank you very much, Mr. Chairman, and I will be happy to answer any questions you may have.

Mr. FUQUA. Thank you.

Yes, we will have questions as soon as we finish with Dr. Buffalano.

Dr. Buffalano, we are pleased to have you here today.

[The biographical sketch of Dr. Buffalano follows:]

Name: Charles Buffalano; mailing address: 1400 Wilson Blvd., Arlington, VA 22209-2308; telephone: 202-694-3035; company or institution: Defense Advanced Research Projects Agency; title: Deputy Director for Research.

As Deputy Director for Research of DARPA, Dr. Buffalano is responsible for the programs conducted in computer science, machine intelligence, communications technology, materials science, and mathematics as well as the demonstration and evaluation of these technologies in military environments.

Dr. Buffalano has been a member of the aerospace technology community for twenty five years. He has served in both technical and managerial positions in the Government and industry at Sikorsky Aircraft, Bell Telephone Laboratories, The BDM Corporation, and the National Aeronautics and Space Administration.

Dr. Buffalano holds a Ph.D. in Engineering and Applied Science from Yale University and a Bachelor's degree from the Stevens Institute of Technology.

STATEMENT OF DR. CHARLES BUFFALANO, DEPUTY DIRECTOR FOR RESEARCH, DEFENSE ADVANCED RESEARCH PROJECTS AGENCY, U.S. DEPARTMENT OF DEFENSE, WASHINGTON, DC

Dr. BUFFALANO. Good morning, Mr. Chairman. Good morning, ladies and gentlemen.

DARPA has a—DARPA is ARPA by the way—DARPA has had quite a major role in the development of computer science in the United States over the last 20 years. Yet, its role is not well understood.

So I would appreciate it if I might take a few moments to tell you what DARPA is and how it operates in this community, and then I would like to discuss the role of the provision of infrastructure to the academic and industrial community, say a few words about networking and the NSF, some words about VLSI prototyping, and make a few comments about the Federal role in software development.

If you think of the Department of Defense as three major operating companies, the Army, the Navy, and the Air Force, you might think the Under Secretary of Defense for Research and Engineering as the chief engineer.

Each of the operating companies, each of the services, has its own laboratory system, peopled by civil servants, and spends basic research money and does research in computer science and extends the boundaries of the computation cycle in the United States.

DARPA works for the chief engineer, and our charter is to do what the services may not do. In fact, DARPA was founded at a time just after the Russian launch of the Sputnik and the Secretary of Defense asked for some scientific activity which would be sure to handle all the things that did not fit easily into other people's charters.

So DARPA is there to protect the Department of Defense and the people of the United States from technological surprise which might be created by a gap in the roles of the services.

DARPA is not a laboratory system; it is not even a laboratory. Each of the services maintains a laboratory, and, as I said, they are peopled by thousands of scientists and engineers.

DARPA is 60 people. The 60 people have a budget of approximately $700 million. And so DARPA really is, the best model you might have of DARPA is, we are a program management agency, or, as the *New York Times* has said recently, we are a group of entrepreneurs. We are the Department of Defense's entrepreneurial activity, and what we will do is like good capitalists: We search out within the United States those people who are doing the best work in areas of interest to the Department of Defense, and fund it.

That has several important aspects for our relationship to the community. First of all, we like to think that DARPA can turn on a dime because we do not maintain a large civil service group. Our laboratories do not have high inertia; we do not have careerists. The average time at DARPA is 4 years.

People who come to DARPA come with a mission. They often are very excited about a particular kind of technology. They are funded adequately, they go outside and find the research and development, they start something and leave.

So we tend to be different from other Federal institutions in that we can move rapidly through science without disturbing the civil service structure.

The other thing that is important about that is that there is a myth that DARPA has a basement and in this basement anything that anybody happens to be interested in at the time is being built by somebody, if only we could find out about it.

There is no basement; there is no such thing. All of the research and development which is done is in fact done in universities and is done in industrial settings so that the technology-transfer problem is better solved because the not-invented-here syndrome does not apply. Again, the downside of the laboratory is that you tend to have people who are the inventors, the parents of technology, and they tend to be covetous.

DARPA is entrepreneurial. We are happy when our ideas find markets and move into the culture of the Department of Defense and the United States as a whole.

For example, I was interested in Al's characterization of the X-29. The forward-sloping aircraft is Grumman's aircraft, and it certainly is, but I will bet the number of people in the room who know that it was funded by DARPA is small.

DARPA tends to do counterculture things. It tends to do things that are unusual.

For example, DARPA developed the AR-15 rifle, which became the M-16 because the Department of the Army did not believe that a 22-caliber round could be lethal. We demonstrated that it could.

DARPA invented stealth technologies at a time when the Air Force was not interested in that.

We have had a long history in computer science from the development of two computers, artificial intelligence, and have provided continuous funding for the major universities who are involved in that today.

DARPA is highly coupled into the civilian community. Because of our venture capital operation, we tend not to deal internally inside the Department of Defense, so roughly half of the 6.1-6.2 money that is spent from DARPA is spent in universities in unclas-

sified work. And so another myth that all this is classified is not the case.

DARPA is a very open organization, very open to the civilian sector. We are members of the President's Federal Coordination Council on Science, Engineering, and Technology, and we are reviewed on a regular basis by the civilians of the Defense Science Board.

Our programs are audited on a line-by-line basis by the House and Senate Armed Services Committees, and particularly our programs in computer science and machine intelligence. The programs are almost entirely unclassified in these areas. What you see and what you read in the popular press and the professional press is indeed what is going on.

One of the important things about DARPA's long-term relationship with the community is that we have been developers of infrastructure. Infrastructure is a buzz word for the sorts of things that we have been talking about today.

Supercomputer access through networking is very important. I was told on a regular basis by people in the computer science community that it is communication which made possible the rapid dissemination and development of their technology, that they tend to work in large communities and the bulletin boards which are maintained on networks are enormously important mechanisms for rapid transfer.

DARPA was a developer initially of the ARPANET, and when people talk about getting on ARPANET, they probably do not realize that it is now managed by the Defense Communications Agency.

So I tend often to be the front for "why can I not get on ARPANET," and the answer is, "you have to go see another fellow," which makes it harder. But we are working with the National Science Foundation now, as a matter of fact, to open additional—to lobby with the DCA for additional tips on networking.

Another what I think is an enormously important common community technology is rapid prototyping. The designers of high-speed machines, both vector and multiprocessors, want to be able to take VLSI designs, very large scale and various circuit designs, and have them checked and tested. They want to be able to design a better machine, get it all on a little chip, and get it back, and try it out, and they want to do that without spending a lot of money. They want to do that as rapidly as possible.

DARPA has developed a system involving the metal oxide semiconductor implementation service, which is run with the University of Southern California, which allows designers to develop designs, and have them turned around, often on boards, and returned within 35 to 50 days.

We are also working with the National Science Foundation to extend that service. In fact, it is currently a shared service with the National Science Foundation to researchers outside the defense community.

I would like to just take a moment to talk about the software role.

DARPA is, for the first time this year, going to fund work in developing generic software for multiprocessors and supercomputers.

When people talk about software, I think it is important to recognize that there are certain kinds of software which are used over and over again. Matrix convertors were discovered once and used 1 million times. The initial value problem is, after all, solved over, and over, and over again.

When we talk about common software, generic software, I think the Government can play a role in developing a library of common routines which are then picked up, and customized, and made specific to a particular industrial user.

In conclusion, those are the three areas in which DARPA is involved in the infrastructure of the supercomputer community, a quick review of how DARPA operates, how it is related to the civilian community, and I hope I have left you with a better picture of DARPA and our role.

[The prepared statement of Dr. Buffalano follows:]

Statement by Dr. Charles Buffalano, Deputy Director for Research, Defense Advanced Research Projects Agency

I would like to thank the Subcommittees for the opportunity to discuss DARPA's present and future plans for supercomputing research and development and network support. I will also discuss how DARPA's programs are coordinated with other federal agency programs.

I will discuss the responsibilities the Agency has in the Department of Defense because it is these responsibilities which shape our policies and programs. I will discuss the relationships the Agency has with the federal community as well as the university and industry communities and speak in particular to our relationships with the manufacturers of computer hardware. Since DARPA and the NSF provide infrastructure and support to the research community I will also discuss our joint interests in networking supercomputers and providing rapid VLSI prototyping. Finally, I will discuss briefly specific areas of DARPA's current R&D interests.

The Defense Advanced Research Projects Agency (DARPA) was created with a very unique and interesting charter in 1958 just after the Russian launch of Sputnik. Niel McElroy, then Secretary of Defense, is quoted as saying that in DARPA he wanted "an Agency that makes sure no important thing remains undone because it doesn't fit somebody's mission." Unlike the Service research organizations which report through the Service Secretaries, DARPA is an independent Agency reporting through the Office of the Secretary of Defense. This allows DARPA to operate independently of the Services to a large extent. However, to balance this independence, DARPA was not given procurement authority and so must work with the Services to get things done. In addition, unlike the Services which maintain a substantial Civil Service laboratory system, DARPA has a total technical staff of 90. And this staff spends an

average time at DARPA of only 4-5 years! These are important facts because they are the basis for the continuous renovation of the technical programs. High turnover means a constant source of new view points and "no laboratories" means that there are minimum pressures to continue technical activities because of institutional commitments. In a recent Wall Street Journal article, DARPA was well described as a venture capital operation, not as a laboratory. Because of these unusual organizational arrangements, DARPA has been able to achieve much over its lifetime.

Because of its charter, interests, and a succession of excellent Directors of the Information Processing Techniques Office, DARPA has played an important role in the development of many of the technologies fundamental to computer science and engineering. DARPA supported the development of computer time sharing, interactive computer graphics, ILLIAC IV, the first high-speed supercomputer, ARPANET, a national communication system for scientists and engineers, packet radio communications, Ada, and artificial intelligence. And since DARPA has no laboratories or scientists of its own, all of this development was carried out in universities and industrial settings. We believe this approach makes technology transfer simpler because it allows local invention.

DARPA is highly coupled into the civilian community. DARPA is a member of the Federal Coordination Council on Science, Engineering and Technology (FCCSET) organized under the President's Office of Science and Technology Policy. This Council meets to exchange information and coordinate research and development activities within the entire Federal Government. In addition, the programs and policies of the Agency are reviewed on a regular basis by

civilians outside the Department through the Defense Science Board. DARPA programs are reviewed on a line by line basis by the House and Senate Armed Services Committees. DARPA programs in computer science and engineering are particularly thoroughly reviewed because of the interests of the members and the staffs and the substantial commercial values attached to these technologies. DARPA programs in computer science and engineering are almost entirely unclassified and so they are also reported and reviewed in the professional and popular press.

First and foremost DARPA has a research and technology development responsibility. DARPA supports the development of critical technologies and provides needed infrastructure to the research community. DARPA is different from some of the other members of FCCSET-DOE, NASA, the NSA and the CIA - which have massive computational problems as part of their operational charters and more like the National Science Foundation and the National Bureau of Standards which provide research support to the operational problem solvers. The National Science Foundation and DARPA currently have joint interests in two infrastructure elements 1) ARPANET for supercomputer access and 2) MOSIS, the metal oxide semiconducter implementation service.

The National Science Foundation, as part of its Supercomputing Initiative, is developing a national network to access new Supercomputer Centers, and support the exchange of information among the supercomputer users. While, initially, the networking activity is focussed on access to supercomputers, it is anticipated that, in the longer term, the NSF Network will form the basis for a general national network to support scientific

research throughout the United States. DARPA will work with NSF to improve that national infrastructure for scientific research.

The network, as currently envisioned, will be based initially on the existing Community networks - e.g. ARPANET, BITNET, CSNET, etc. In addition, the NSF is planning a series of pilot projects to explore the use of 56KB and higher terrestrial and satellite links for access from terminals and sophisticated workstations to remote Supercomputer Centers, as well as to provide input to the design of the future NSF Network. NSF and DARPA plan that a number of supercomputer and user sites will be connected to the ARPANET and that the ARPANET may be used by NSF supercomputer grantees to access the remote Supercomputer Centers, and to exchange information.

MOSIS is a service which permits researchers and designers across the country to rapidly acquire prototype custom-designed integrated circuits and printed circuit boards. MOSIS provides access to state-of-the-art fabrication lines at low cost through cost sharing using multiproject wafers. Circuits specifications are transmitted via the ARPANET communications network to the MOSIS Center at USC-ISI, which coordinates manufacture, test and product delivery. Manufacturing times from receipt of data at MOSIS to shipment of assembled parts in as little as four weeks have been obtained. The availability of the MOSIS service to the research community provides a unique capability to implement new ideas in VLSI design techniques as well as novel system architectures.

Our relationship with computer manufacturers are diverse. Since DARPA does not have a computationally intensive operational charter, it is not perceived

directly as a major market for computers. However, DARPA is important to the computer manufacturers, because DARPA's research community does need computers and they obtain these under their DARPA contracts. This is particularly important to the research community in the universities because it is a major source for them of computers and particularly of advanced machine architectures on which to do research and train students. Each researcher decides what machines are needed and there probably is a DARPA researcher on every known machine. DARPA also buys time for its researchers at other Government and industrial facilities.

The major impact DARPA has, of course, on computer manufacturers is through its research program. DARPA and the computer manufacturers have joint interests in basic computer engineering technologies 1) optoelectronic interconnects 2) high density RAM 3) wafer scale integration and 4) digital gallium arsenide integrated circuits. DARPA and the machine architecture community are sharing the development of coarse and fine grain machine architectures which we expect will provide the ability to compute beyond the gigaflop region. DARPA is also interested in 1) the development of algorithms which will exploit multiprocessor architectures and 2) the discovery of important physics at the limit of computing.

DARPA understands and recognizes the importance of supercomputers and has an important role to play in the development of the technology base which will support this new and wonderful class of computing machines.

PANEL 1 DISCUSSION

Mr. FUQUA. Thank you very much. We appreciate your remarks. We will bend the rules of the House of Representatives only because this person has served as a member of our committee in the last Congress, and the good people of Tennessee elected him as their U.S. Senator, and he is going to be the luncheon speaker today, and we are very glad to welcome Senator Albert Gore from Tennessee.

Normally, we do not do this, Al, but since you are a former member of the committee, we will make an exception only for you.

Senator GORE. Well as a former member of the committee, I fully understand what an exception this is. Thank you.

Mr. LUJAN. You need to balance off the other side of the aisle.

Mr. FUQUA. Dr. Trivelpiece, you mentioned in your testimony that 37 percent of the fiscal year 1985 supercomputing hours were allocated to university users. What policies in DOE govern the allocation of this time versus the universities—or the universities versus say the national laboratories?

Dr. TRIVELPIECE. By response to requests, the constraints with which we operate is that we cannot make time available to other than DOE contractors, and so from within the university-contractor-grant community that the Department supports as well as the laboratories, we receive proposals for the use of time, which has been set aside on the MFE network. Those proposals are peer reviewed and they are selected on the basis of their quality. And I suspect that the 37 percent is a consequence of that competition and is not in the least part of the policy stated at the outset. I think it is the outcome rather than a policy statement.

Mr. FUQUA. Do you think that will change in fiscal year 1986?

Dr. TRIVELPIECE. I do not know. If it is in response to, then I think the consequence is it will depend on the degree to which the university community versus the laboratory community prepares proposals that compete.

Mr. FUQUA. But is there not some distinction between computers which are used for numerical calculations and those that are used for artificial intelligence application? What are the differences and similarities between the two?

Dr. TRIVELPIECE. The differences exist at three levels as I see it. Two of them are social. One appears to be technical. Computing machinery all the way down on the inside starts to look awfully alike. You can probably tell a numerical machine by going to find out where the arithmetic units are and where the floating point system is. And if you looked way down deep inside and did not find that, but you found a lot of little wires connecting up all the processors, you would probably think it was a semantic network.

I think what we discovered in the last year in the development of architecture is that a machine that is good for one often tends to be good for another, and that adding a floating point operation to a symbolic machine does not cause major repercussions.

So I think the differences are going to disappear over time and I would hope the distinctions eventually will as well.

The other two tend to be social; that is, the communities are separate. The artificial intelligence community and the computer sci-

ence community have not fully integrated and so they tend to like the devices they are comfortable with. LISP is a language which the community has adopted in AI, and so if a machine runs LISP, it tends to be a symbolic processor. So I think that those are the kinds of differences; at a technical level I think they are minimal and at the social level they are not minimal.

Mr. FUQUA. Dr. Good, in the NSF plans for fiscal year 1986, are you planning to continue to support the computational centers that have already been named, or do you plan to allocate funds for other activities in addition to the national centers?

Dr. GOOD. Mr. Chairman, with funds available, we hope to do both. We do plan to continue to support those centers that are there. They have been extraordinarily well used and it really would not make a lot of sense not to continue support for those.

How much further we can go and how many new things we can add in 1986 obviously at this point in time depends on budget allocations and what kind of movement within the budget we can make if we end up with a level budget.

Mr. FUQUA. And also the operational cost of those four centers.

Dr. GOOD. That is correct, and the operational costs of those was reasonably well worked out when we made the original plans. One of the things we do try to do is not get caught later with a center that comes back and says "Well, yes, I told you I could operate with this kind of money, but really to be effective I am going to have to have twice that kind of allocation."

We make every effort not to get caught in that kind of situation, so I think that has worked out reasonably well.

Mr. FUQUA. Thank you. Mr. Lujan.

Mr. LUJAN. Thank you, Mr. Chairman.

Dr. Trivelpiece, I guess I misunderstood you. I thought in your answer to Chairman Fuqua about university use of DOE facilities, computer facilities, that only those universities that are contractors with DOE—does that mean a place like Livermore can only allow the University of California who runs it, access to the supercomputer?

Dr. TRIVELPIECE. No, not just the University of California, but any principal investigator at any university who is a contractor with the Department of Energy can then, if he submits a proposal for access to or time on and have that awarded, but it is on a competitive basis within the time that is allocated for that branch of science; for instance, high energy physics, the 5,000 hours for high energy physics do not just simply go to the laboratories. It could go to a university as well.

Mr. LUJAN. But that is work that would be related to a DOE contract that they might have?

Dr. TRIVELPIECE. That is correct.

Mr. LUJAN. Not just some other work that they want to do.

Dr. TRIVELPIECE. That is correct.

Mr. LUJAN. Dr. Good, in talking with different people about the centers, they mean different things to different people. Some see it as making Tallahassee and this part of Florida a new silicon valley. Some see it as a great opportunity to try out their new engineering projects. Others see it as an educational thing for the univeristy.

How do you see it, as a study center or for industry, for researchers?

Dr. GOOD. That is a good question and obviously the comments you have heard are very true. Frankly, my personal hope is that they are perceived in all of those ways for a very unique reason. The National Science Foundation does not have in its budget, either now or projected, enough funding to support the supercomputer needs of everybody in the country, and if we can get the local people who are in those centers very excited and if we can get them to talk about the kinds of utility that those centers can have all the way from the educational aspects, the technical aspects, getting industry involved in those centers; the more we can do that, the more we can leverage the NSF investment in those centers.

And frankly, that is really what we have to try to do because I think once again the NSF money is not going to be able to support all of those centers in their totality, or will we be able to have the funds available to do all of the things those people want to do.

But if we can get them to believe that they are all of the things that you just said, then there are lots of other places that you can get in the game, and that in a sense is really what the NSF is trying to do, trying to leverage what funding we have as much as we possibly can, getting industry interested and letting them pick up part of the cost because they use it, and getting the educational establishment involved, getting the State deciding the educational issue and putting up some funding.

So I am hoping that they are indeed perceived as all of those things and will indeed provide for all of those things.

Mr. LUJAN. So you think each one of the centers will be different and it is for each center to provide their own program of what they want to give priority to. I thought that at least in the case of one member of the committee, myself, that they were primarily educational tools rather than—well, you know, I have a particular problem, but I will not buy a computer or I will not buy some time. I will just go on over to the university and let the boys do it for me.

Dr. GOOD. Well, I think that that latter possibility really is not there any more, because you cannot have something for free. Computer time was at one time very much looked upon as a library is looked upon. There is a certain thing which says that libraries are free, right? They have always been there and all the information is there and available to you, and all you have to do is come in the front door and get it.

Well, frankly even that is not true any more because the kinds of information that is coming to the library is coming in on electronic devices and somebody has to pay for that. So I think you are going to see the same kind of thing at the centers.

Now, the NSF support part, part of this will come from the NSF and obviously there are two major components of that or the NSF would not feel it appropriate to fund that. The first is that we feel it must make some impact on science, that is what we are all about. If it does not allow you to do new scientific problems or do the problems you are presently doing better—that is one of the necessities. And you have to have people who are associated with it who are competent to do that.

And the second one is the education issue. It must be utilized for the development of new people who are able to use these tools at the forefront.

Those are the two major issues on which the NSF makes the decision as to where the funding goes. Now, beyond that, we would like to encourage them though to do all of these other things. It allows them to leverage that center and get a critical mass that is bigger than we could just with those two issues.

And frankly the other point is that particularly the interaction with the industry, the educational aspects of that are very high. If you get a lot of industry problems in there and there are graduate students mixing with those people, the educational aspects are highly improved.

Mr. LUJAN. Thank you, Mr. Chairman.

Mr. FUQUA. Mr. Sensenbrenner.

Mr. SENSENBRENNER. Yes, I would like to follow up on Mr. Lujan's line of questioning.

It is obvious that everybody would like to get a supercomputer center because it is the newest and niftiest thing on the block. And I think there is a great danger that the congressional pork barrel will replace scientific merit in terms of allocation of the scarce dollars. What do the DOE and the NSF plan to do to stop that, recognizing that we hold the purse strings?

Dr. TRIVELPIECE. He asked you first. [Laughter.]

We try to receive proposals and have them evaluated on the basis of merit, and the guidance that is provided to us from time to time from the Congress. I think this is a matter for you to review and decide how it is best done.

We have made our plans and we make our plans known to you through the President's budget which is submitted each year, and then we appear before you and debate these matters or present our case to you. But to some extent, you have a rather substantial amount of control over this activity. I am not sure what else we can do beyond that which we have already done to see to it that the playing field is maintained as flat as possible and fairly managed in the operation. We are dedicated to doing that.

Dr. GOOD. The National Science Foundation's review of these types of proposals are based on exactly the same criteria that we use on all other types of programs. The Foundation requires that the people who wish to participate have to put in proposals and there is a program announcement so that everyone who has an interest in it has the same shot at putting their proposal together.

The decisions—and I can assure you that all of the ones that I have discussed today—the decisions on which one is to fund are absolutely based on those two things I just said—the best science and the best scientists attached to that, and the best program for how they are going to utilize that and how they are going to leverage these educational issues. Really, those are the criteria and I think in the sense that we adhere to those criteria, the problems that you discuss will go away.

Now, the real question comes down to the crunch of what people can afford and what they need. There is a great pressure today—and I think it is fair to say that it is not without merit—there is a great pressure out there because an awful lot of universities in the

country today are in very difficult trouble in terms of trying to maintain first rate research programs and things of that sort, and some of them are in the process of looking for ways to get what they need by whatever mechanism is available to them.

I am not sure I am opposed to that. It is just that the NSF program will have to adhere to the standards of quality and if we do that, then I think we can handle the situation reasonably well.

Mr. SENSENBRENNER. Thank you.

Mr. FUQUA. Mr. Boehlert.

Mr. BOEHLERT. Dr. Buffalano, what level of supercomputer support is DARPA providing the universities?

Dr. BUFFALANO. If within supercomputer, you—you will permit multiprocessors, which, of course, Cray-XMP is a multiprocessor, we are currently funding though the strategic computer program, the basic program for machine intelligence, on the order of $80 million in just hardware architecture and software development for those machines.

Mr. BOEHLERT. Did you say $80 million? And how about industry, how much to industry?

Dr. BUFFALANO. That $80 million is split 50/50 between industry and the universities.

Dr. BOEHLERT. Dr. Good, where would you like to see us 10 years from now? We have just had these four centers announced and we are looking toward the future and trying to get other people excited about the potential, where would you like us to be, how many centers?

Dr. GOOD. In many ways, the total number of centers I think is really not the issue. The issue is to have access.

Mr. BOEHLERT. Networking?

Dr. GOOD. Networking. I would think that really the next piece—with the establishment of the centers that we presently have—that the next real goal ought to be for the networking. We need to find out what the real demand is out there and we need to know how many people truly do need supercomputers.

I have a little bit of a problem. It is true that we need to have these, we need to have access to them, but on the other hand, there is an awful lot of work that needs to be done and many really first class scientific problems today really do not need supercomputers, they simply need access to small minis and things of that sort.

So I think we need to look at the whole computing infrastructure in addition to just the centers. The whole concept of networking seems to be really the next step. How do we do that, and how to get the really first class research person who needs access, who is in a university where he may be the only one of two people who need that, and how to provide that access to him at a cost we can afford to support.

Mr. BOEHLERT. Would you suggest then if we are dealing with the limited resources we have today, that rather than thinking in terms of new centers, we put more into the existing centers that are already underway?

Dr. GOOD. Surely that is one of the things that should be done, because one of the other tracks that is always true in one of these initiatives is that if you continue to proliferate the number of cen-

ters, then you will within the 10 year period you have just described, at least the front end of those centers will be obsolescent.

I think the real problem will be to keep those centers at the cutting edge. If we are going to put that kind of funding into them, we ought to see that they are adequately supported, and then make access to them available to a much larger group of people.

I think as you understand the whole concept of networking over the next 10 years, I think it is going to change dramatically. The quality of networking is going to improve. Even our best networks today are not necessarily state-of-the art. So one of the things we need to do is to upgrade the quality of the networks we have, looking at the technology that is available with satellites, the other kinds of things, and improve the current networking. And I would like very much to see us put a fair amount of support into that particular activity.

Mr. BOEHLERT. If you watched television this morning and the program "Good Morning," you saw our distinguished chairman addressing that very same subject.

Mr. Chairman, I also would like to recognize the presence of our colleague from the other body, Senator Gore, and note for the record that in this last election, three Members of the House of Representatives opted for the longer term life over in the Senate, and all three of those came from the Science and Technology Committee.

Mr. FUQUA. Mr. Boucher.

Mr. BOUCHER. Thank you, Mr. Chairman.

Dr. Buffalano, it is my understanding that spending about $50 million annually, DARPA is conducting the lion's share of federally financed research into artificial intelligence. I am interested in knowing the extent to which the results of that research will be classified and to the extent that the results are not classified. What plans does DARPA have for transferring the technology which you will create to the private and to the academic sectors?

Dr. BUFFALANO. Currently none of the programs on artificial intelligence is classified. Also, all of the program that is currently being supported in artificial intelligence is being done in universities and in startup venture companies and in industrial settings, and everyone is free to publish and to share that information. That is a special feature that you get when DARPA runs the program. That is an important thing and the transfer happens automatically.

We are supporting, for example, basic research in AI in startup venture companies. They will sell it. That is not the goal of the Department of Defense to create a commercial market, but it will happen.

Mr. BOUCHER. That is very encouraging to hear. So DARPA has no plans to recommend classification for any of the technology which will flow from the AI research?

Dr. BUFFALANO. The 6.1 and basic 6.2 research are at this point not classified at all, and there are no plans to do so.

Mr. BOUCHER. Dr. Good, you had mentioned the work that the NSF is doing in promoting the national network for supercomputer. Can you give us some indication of when you think a nation-

al network created through the efforts of NSF will be up and running?

Dr. Good. We would like to have it up and running in about 3 years time. Whether or not that is an achievable goal is a separate issue, but we would really very much like to have that. That is the time frame we would like to see it built.

In the interim—because we cannot wait to get that done—in the interim, and as I said at the June meeting we will discuss, as Dr. Buffalano has said, we will discuss the access through ARPANET for a fairly large number of increased notes. So this is going to help a great deal. We are going to look at things like the BITNET, which is not as satisfactory as the ARPANET, it is a little bit slower, has some other constraints, but on the other hand for people who have no access at all, BITNET is quite good.

So, you know, we are going to try at the front end to do as much as we can with the existing systems and at the same time, look at the potential of really upgrading the whole quality of networks and looking at a network in place, and we would hope to have that in about 3 years. Now that is very ambitious, but we will try.

Mr. Boucher. Do you feel that the funds that you have for the current fiscal year, as well as your anticipated budget for fiscal year 1986, will be sufficient to allow you to bring a national network on line within a 3 year period?

Dr. Good. We surely do at the present time, unless we have some unpleasant surprises. We think the answer to that is yes. And in fact, we have planned within that constraint to try to do that.

Mr. Boehlert. That is excellent news.

Let me ask a question for all three of you and that is this: in the very broad sense, where does the United States stand in this supercomputing effort in relation to other countries, primarily Japan, that also has major initiatives in this area?

Dr. Trivelpiece. That question is probably better addressed to some of the technical experts that are going to follow on the next panel, as well as it will be a subject of the conference. Certainly the Japanese have prepared and designed and developed some advanced machines which are competitive in every practical sense.

The vendors are also I think in a position to comment on that, but I do not think the United States has given up its lead in any practical sense yet and I certainly hope that it does not do so. As I said in my opening remarks, I believe it is essential for the long term economic health of the United States and we should make every effort to see to it that we do not lose that lead that we do have.

Dr. Good. I again am in the same position. I cannot really comment on the technical position, but I would say that the advent of the networks, the kind of things that are going on at DOE, the things that are going on at DARPA, as well as the NSF programs, provide something that the Japanese do not have and that is diversity. They do provide a large group of people doing, if you will, their thing, and frankly, in the past when we have been able to do that, we have been able to compete at whatever level we like, and I think that really is the significance of the different centers for different things.

And I think the worst thing you could possibly do is to have a monolithic approach to that problem. The diversity and the ability to have people sort of set loose to do their thing, in my opinion, is the best position you can have for competition. With the access that has been created over the last 3 or 4 years, I think we will be able to do that. I am not very pessimistic.

Mr. BOUCHER. Dr. Buffalano.

Dr. BUFFALANO. I am positively optimistic. Let me suggest that you might look at the President's Foreign Intelligence Advisory Board review of this very question and whether or not the sum of the activities within the Federal Government are adequate to maintain that lead. I would like not to tell you what it says at this point because it has not been formally released and I do not want to preempt the chairman, but I think you will be surprised and pleased.

I would also like to say it is clear and that with respect to the Russians, that the United States has a commanding lead in any technology at this point related to computer science. We have a 10 to 15 year lead in hardware and software. I have to say that is not a classified number.

So anything we do in the United States to press that advantage has enormous value and the speed with which we move these technologies into the military systems of course makes that lead even better.

Now, in this committee, of course, I know you are interested in national defense, but on the other hand, it is like a commercial spinoff too, which finds its way almost as rapidly, and in some cases as rapidly, in the civil sector.

The Japanese Program, as you know, has three major parts: one is fifth generation, which is artificial intelligence, machine intelligence program. I think that our programs are as good as, if not better, than the Japanese Program and we will succeed.

Mr. BOUCHER. Dr. Buffalano, could I interrupt you on that point for just a second? How does your budget, being the bulk of the federally funded AI research, compare to the ICOT budget in Japan, which I understand is spending the bulk of their government-financed research?

Dr. BUFFALANO. At the current level, ours is roughly twice the Japanese last year's spending, which will make it roughly the same this coming year. But the important fact is that it is not the bulk of the spending that is being spent in Japan. There is very important work in artificial intelligence going on in the corporations. ICOT is, after all, a consortium activity, which is composed of some 40 people or so who are working on, who are borrowing people from the industries, training them, building prologue machines, and sending them back.

So, what you see at ICOT is only the top of that activity, and I have no way of knowing what the scale of the activity in artificial intelligence is within the corporations. But I do know that it is substantially larger.

Mr. BOUCHER. Do you feel we are being competitive with the Japanese in AI research today?

Dr. BUFFALANO. Yes.

Mr. BOEHLERT. Would my colleague yield?

Mr. BOUCHER. Yes; I will be glad to yield.

Mr. BOEHLERT. Are any of you satisified—and let us go right back to the beginning—with what we are doing at the elementary and secondary level? I mean, should we not all be alarmed that we are not doing nearly enough in terms of preparing our young people in computer literacy and that type of activity? I read just recently where the Soviet Union has embarked on a program to have PC's for all their kids in high school and the Japanese certainly are doing——

Dr. BUFFALANO. Do you know where they are buying them?

Mr. BOEHLERT. No.

Dr. BUFFALANO. They are going to buy them probably in the United States because they cannot build adequate——

Mr. BOEHLERT. That is fine for our balance of trade and all that sort of thing, but the basic point is that every youngster coming out of a Soviet high school is going to be a computer literate, and the same thing in Japan and other nations, and we are not doing nearly enough in that regard. Should we not be focusing more attention on that?

Dr. BUFFALANO. Certainly. How can you not want to be for education?

Mr. BOEHLERT. Dr. Good, how about it? We are all for it, but we are not doing a heck of a lot. That is a concern I have.

Dr. GOOD. That is true. However, there was a new scorecard put together last week, I think, which was publicized, on what the States have done with respect to education, where the upgrades are happening, and although they are not dramatic in most cases, if you look at the overall scorecard, we really have done more perhaps than we think we have in terms of improving the curriculum requirements and the upgrading of the instructors.

One of the biggest problems of doing the kind of thing that you are suggesting in the elementary schools today, is the possibility of finding young people who could teach that and who have the facilities and the resources to do it.

We do not have a cadre of elementary school teachers who could attack it, so somehow you have got to—it seems to me you have got to do that first, because to teach it poorly may very well be worse than not teaching it at all. And this is one of the problems that we have.

But it is true that I think we are getting ourselves to a point where we have a lot of people who—a small percentage of students who have access to lots of things and we have a very small group who really are at the active forefront of everything—we have a group of students who are in that category, but there is getting to be a very large gap between that group of maybe 5 or 10 percent, and the rest of the students.

And what we should try to do something about that, is hard to say, but we are going to be hard-pressed in the industry to do the kinds of things you want to do, in just hiring technicians and things of that sort, never mind the scientific component.

Mr. BOEHLERT. Thank you.

Mr. BOUCHER. I believe my time has expired, Mr. Chairman.

Mr. FUQUA. Senator Gore.

Mr. GORE. Thank you, Mr. Chairman.

I would like to begin by saying I have worked in this area for awhile, and I do not know of any conference or hearing or gathering anyplace in the world that has brought together the kind of talent that is assembled here in Tallahassee today. If there is one leading expert in this field who is not here, I do not know who in the world it would be. And I am just glad to have a chance to participate.

I only have a couple of brief questions. First of all, Dr. Good, you have talked about the network linking together supercomputers. What is that network going to be made out of, fiber optics?

Dr. GOOD. That is a very good question. I had breakfast with John Connolly this morning to discuss that issue with him, and I believe our technical people—that is a good question. The possibility of it being fiber optics is very strong, but all the votes are not in yet and we are just going to have to take a look at it, but surely the possibility I think of fiber optics is a very high one.

Mr. GORE. Well now, is it not reasonable to expect that over the next—let us look not just at ten years out, but 25 or 30 years out—is it not reasonable to expect that there will be an ever-increasing need for expanding that fiber optics network as more such machines are located throughout the country?

Dr. GOOD. Surely, no question about that.

Mr. GORE. And who is going to pay for the fiber optics network?

Dr. GOOD. Well, Mr. Trivelpiece says I should tell you AT&T. I do not really believe that.

Mr. GORE. Well, I had a meeting with AT&T about that a couple of years ago, and they have an ambitious project but it is still very small in scope compared to what ought to be contemplated to link these machines up around the country.

And the reason I bring this up, Mr. Chairman, is to just propose an idea for consideration by your committee and by the Senate. It is something that would take a great deal of time to think through and work out, a great deal of expense obviously would be involved. It is something to just begin considering.

In the 1950's when this country wanted to dramatically increase its industrial power and increase its economic strength, we had a national program to build the interstate highways. In this new era we are now entering, it may well be that toward the end of this century—I would like to see it even sooner than that—but it may well be that we ought to consider a national initiative to build interstate highways for information with a fiber optics network connecting the major computational centers in this country, with provisions to allow access to those with lesser machines, and to consciously fuel inject to the Nation's ability to take the best advantage of this new technology.

Let me ask one other question.

Dr. Buffalano, in response to Congressman Boucher, you confined your remarks to artificial intelligence. He asked about the classification of the research and I take it that the premise of the question is similar to an idea that has been bandied about quite a bit, and that is that the current inability of the research community in the United States to really make good use of supercomputers is in large part due to the fact that they were developed and first used in a compartmentalized way and people in the research community

are not really familiar with them. They have not had access to them, et cetera, et cetera.

There is kind of a worry that the Strategic Defense Initiative and the computing component of it is going to push the frontiers of knowledge back very quickly toward sixth generation computers and yet we are going to have the same problem in making use of it for the other things in the country besides defense. Now, is it not true that outside of AI, there are large chunks of this that are going to be strictly compartmentalized?

Dr. BUFFALANO. There are several things in the question which I would disagree with as a model, but——

Mr. GORE. Well, please challenge me. I am here to learn.

Dr. BUFFALANO. As I recall, the ILLIAC IV was one of the first multiprocessors and it was at NASA Ames and open to the public. As I recall, the high speed vector processors were not created in compartments.

Paul, is that not correct?

[Paul Sneck, the expert, is nodding.]

Mr. GORE. Maybe the DOE——

Dr. BUFFALANO. I am not sure that that model, that they were created in compartments, and then sprung on the community is correct.

But let me answer the question too, which you asked, which is what about strategic defense? Are they going to push the state-of-the-art and do it in compartments? Well, let us unravel two things: One is DARPA does not run the Strategic Defense Initiative program. That is not quite the answer to the question yet, but the Strategic Defense Initiative is, as you probably know, a separate agency.

Mr. GORE. Yes, I know all about that, but the computing part of it——

Dr. BUFFALANO. They have a great dependence at this point in time on the hardware architecture developments and on software developments of DARPA. And until they really begin to build architectures out of the bag of architectures which has not yet appeared, since DARPA and the Navy are funding many—it does not seem to me that there will be anything classified.

Mr. GORE. Well, I will explore that with the next panel, too.

Thank you.

Mr. FUQUA. Thank you very much.

We thank all of you for being here this morning.

[Questions and answers for the record follow:]

POST-HEARING QUESTIONS AND ANSWERS

Relating to the

FIELD HEARING - TALLAHASSE, FLORIDA

on

JUNE 10, 1985

Before the

SUBCOMMITTEE ON ENERGY DEVELOPMENT AND APPLICATIONS
and the
SUBCOMMITTEE ON SCIENCE, RESEARCH AND TECHNOLOGY
of the
COMMITTEE ON SCIENCE AND TECHNOLOGY
HOUSE OF REPRESENTATIVES

WITNESS: ALVIN TRIVELPIECE
DIRECTOR OF THE OFFICE OF ENERGY RESEARCH
DEPARTMENT OF ENERGY

Question 1: What coordination is taking place among the DOE, NSF, and DARPA supercomputer programs? What other federal agencies are involved in this area?

Answer: In 1983, the Office of Science and Technology Policy established three subcommittees under the Federal Coordinating Council on Science, Engineering, and Technology to examine the U. S. role in maintaining leadership in the development and use of supercomputer technology. Membership on these subcommittees includes representatives from the Department of Commerce, Department of Defense (including the Defense Advanced Research Projects Agency, Central Intelligence Agency and National Security Agency), Department of Energy, National Aeronautics and Space Administration, National Science Foundation, and the Office of Science and Technology Policy. The subcommittees were charged with developing recommendations on the role of the government in maintaining the U. S. lead in supercomputers, on providing access to supercomputers for our researchers, and on coordinating the roles of each agency in the area of research on artificial intelligence and high performance computing. Each of these subcommittees produced reports for the Office of Science and Technology Policy and continues to provide coordination among agencies in these areas.

Two of the subcommittees, one addressing government supercomputer procurement policies and the other addressing access to supercomputers for researchers, were combined into a single subcommittee, chaired by Dr. James Decker, Department of Energy. This subcommittee maintains contact with industry representatives, sponsors panels to investigate questions concerning computer network access to supercomputers and interagency sharing of resources.

The third Federal Coordinating Council on Science, Engineering, and Technology subcommittee is chaired by the Department of Defense and its charter is to stimulate the exchange of information within the government on high performance symbolic computing and artificial intelligence. During the course of its initial deliberations it became clear that interest in information exchange extended beyond symbolic processing and artificial intelligence. It was decided to broaden the charter of this subcommittee to include federal research efforts in very high performance scientific and numerical computing and advanced computer research. A report, entitled "Report of the Federal Coordinating Council on Science, Engineering, and Technology Panel on Advanced Computer Research in the Federal Government," June 1985, has been submitted to the Office of Science and Technology Policy for further action.

Another interagency group that has been coordinating research for many years is the Interagency Committee on Extramural Mathematics Programs. Active members of this committee include representatives from the Department of Energy, Department of Defense (Air Force Office of Scientific Research, Army Research Office, and Office of Naval Research), and the National Science Foundation. Occasional participation in the Interagency Committee on Extramural Mathematics Programs activities include representatives from the National Institutes of Health, National Aeronautics and Space Administration, Central Intelligence Agency, National Security Agency, and the Department of Commerce (National Bureau of Standards). In FY 1985, the Interagency Committee on Extramural Mathematics Programs cooperated in establishing the Board on Mathematical Sciences in the Commission on Physical and Mathematical Sciences of the National Research Council in the National Academy of Sciences.

Question 2: What are the critical areas of supercomputer R&D?

Answer: There are two broad critical areas of supercomputer R&D: software related issues and hardware related issues.

In the software area, new problem specifications and algorithms for implementing these specifications are being studied to take advantage of the parallelism offered by new supercomputer architectures entailing multiple independent processors working on a single, large problem in order to achieve the increase in capability required for realistic models of physical processes. These new architectures, currently having up to four processors, are expected to expand in the near future to 16 processor systems. In the longer term, hundreds or even thousands of processors will be coupled into a single powerful supercomputer system.

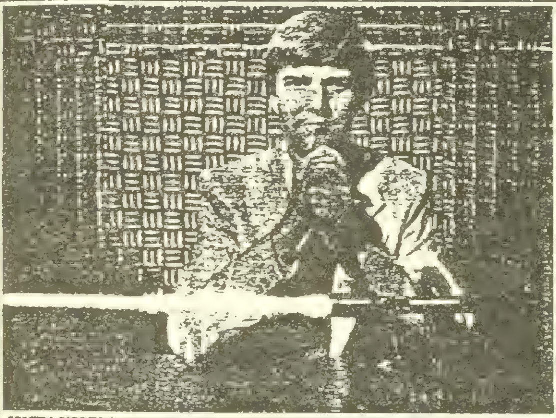
In the hardware area, new component technologies, such as gallium arsenide and high-electron-mobility materials are under investigation. As an example of DOE funded advanced materials research pertinent to supercomputer development see the attached article entitled "Lifting the Speed Limit on Chips" from the August 5, issue of Business Week.

Science & Technology

RESEARCH

LIFTING THE SPEED LIMIT ON CHIPS

SUPERLATTICES—100 TIMES FASTER THAN SILICON—COULD HELP OPEN UP A NEW ERA IN COMPUTERS



SPECTRA DIODE'S SCIFRES: HIS LASER IS THE FIRST PRODUCT TO USE A SUPERLATTICE CHIP

Scientists trying to build faster computer chips are finding that they have a lot in common with highway engineers. Today's silicon chips are like a one-lane road. No electron can go faster than the one in front of it. So to speed things up, they are building unique man-made materials that provide passing lanes for the fastest electrons.

The trick is a new generation of semiconductors known as superlattices. These combine extremely thin layers of different semiconductor materials into a single multilayer chip. Their unique crystal structure enables fast-moving electrons to "jump" from one material to another. Theoretically electrons will be able to zip through the tiny circuits 100 times faster than in silicon chips and 10 times faster than in gallium arsenide, a costly semiconductor material that is beginning to be used where speed is a priority. Already, scientists have built semiconductor chips that operate ten times faster than silicon and four times faster than gallium arsenide.

The promise of superlattices is "extremely dramatic," says Frederick L. Vook, director of solid-state sciences at Sandia National Laboratories. "They will change semiconductor electronics in a big way." Indeed, superlattices are under development at more than a dozen corporate and university labs including IBM, AT&T, and the University of Illinois. Japan's Fujitsu, Hitachi, and Sumitomo—involved in their country's effort to build superfast computers—are also pursuing the technology.

AN ATOM AT A TIME. Speed is only part of the promise. The combinations of materials permit semiconductor scientists for the first time to vary the properties of semiconductors. They can pick and choose among pairs of semiconductor materials such as gallium arsenide and gallium aluminum arsenide or germanium and silicon.

And they can control such semiconductor properties as optical characteristics and operating temperatures by varying the amounts of these materials. "The most exciting part of this work is that you can really engineer absolutely new materials which you just cannot find in nature," says Venky Narayanamurti, director of solid-state electronics research at AT&T Bell Labs.

The new chips are laboriously assembled, almost an atom at a time, by costly crystal-producing techniques such as molecular-beam epitaxy. This process uses a beam of electrons to deposit layers of atoms so thin that 40,000 of them would equal the thickness of a sheet of aluminum foil. Some chips have as many as 40 layers of alternating materials.

Despite the high cost and still experimental status of most projects, superlattices are already coming out of the laboratory. The first commercial product is a laser that relies on several superlattice chips, each the size of a grain of sand. "It packs up to 10 times the power of a silicon semiconductor laser of its type, yet runs on the energy it takes to power a flashlight," says Donald R. Scifres, president of Spectra Diode Laboratories Inc., a joint venture of Xerox Corp. and Spectra-Physics Inc. in Mountain View, Calif. Several government labs and military contractors already use the device in satellite communications links.

Other superlattice devices are in the pipeline at a number of research centers. AT&T Bell Labs, for instance, has experimented with superlattices in a wide range of applications, including transistors capable of ultra-high-speed switching, low-current lasers, and sensitive light sensors for use in fiber-optic and satellite communications systems.

STRETCHING CRYSTALS. Just last fall, scientists at Sandia developed a new technique for making superlattices that permits them to combine materials that previously were incompatible. The laboratory has already demonstrated a laser that uses the new "strained-layer" superlattice. And it has filed for patents on the laser along with a dozen other superlattice devices, including a highly efficient solar cell.

However, it is still too early in the development of this new class of semiconductors to attract the major chipmakers such as Intel, National Semiconductor, and Motorola. In fact, most experts believe that it will be a number of years before the impact of superlattices is felt in computers, where low cost is a priority.

But the big semiconductor companies are all watching the developments with intense interest. "We are now in superlattice technology where silicon was in the late 1940s," says George H. Heilmeier, chief technical officer at Texas Instruments Inc. And they all admit that superlattices could prove to be a key technology for building the next generation of superfast computers.

By Claudia Ricci in New York

ED KASHI

Question 3: What is the private sector doing in supercomputer R&D, and how are the federal programs coupled into these private sector activities?

Answer: The supercomputer vendors have substantial research and development programs specialized to their particular interests. Details of much of this research are considered proprietary. However, the industry has shown a willingness to cooperate with university and laboratory research programs by providing early access to, or substantial discounts on, prototype production machines such as the single quadrant Cray-2 that was made available to the Department of Energy at the National Magnetic Fusion Energy Computer Center, the cooperative program with Florida State University supported by Control Data and ETA Systems, etc.

The Federal Coordinating Council on Science, Engineering, and Technology subcommittee on supercomputer access provides a forum for government and industry information exchange and cooperation.

Question 4: We hear the possibility of having a "supercomputer on a chip" that may have the computing power of today's large machines at a fraction of the cost. Is this a "pie in the sky" notion, or is it a real possibility, and if it is real, how soon might we see such devices?

Answer: Recent announcements have made the trade press in the form of "superminicomputers" on a chip (such as the DEC Microvax II and the Data General MV12000, which is actually on five chips). These superminis are performing in the 5-10 millions of instructions per second (Mips) range.

Also coming on the market are the new "minisupercomputers," such as the Convex 1, which is advertised as "one-fourth the capability of a Cray-1 at one-tenth the cost." To keep things in perspective, however, note that the Cray-1 is now almost ten years old and perhaps should no longer be considered a supercomputer. The recently delivered four processor Cray-2 has more than ten times the capability of the Cray-1 at approximately twice the cost of the Cray-1. Thus, the cost/performance relationship between the new "minisupercomputers" and the new Class VII supercomputers in 1985 is roughly one thirty-second the capacity at one-twentieth of the cost.

As the feature size of integrated circuits continues to be reduced, it will be possible to put more and more logic circuits and memory on a single chip. However, it seems likely that computers meeting the definition of a supercomputer, i.e., the most powerful machine available for scientific calculations, will still employ a number of these chips as processors with local memory. The total supercomputer system might well use many such chips plus a large shared memory.

Question 5: How are hours of supercomputer time allocated? What policies have been established in the allocation of time among different disciplines and among researchers within a given discipline?

Answer: The supercomputer access program within the Office of Energy Research (ER) is managed in the same manner as other research activities. The resources are allocated throughout the research community on the basis of scientific merit and programmatic need. These resources are not allocated in units of hours but rather in units of Computer Resource Units (CRU's) which take into account other resource factors, such as memory usage, I/O access, file storage usage, etc., in addition to central processor unit (cpu) usage. Allocation of resources through this program is restricted to ER contractors, therefore, at the beginning of each fiscal year, requests for supercomputer resources are submitted to the ER program offices at DOE Headquarters in order to verify that requesters are, in fact, current ER contractors. These requests are submitted in response to a "Call" or solicitation that explains the evaluation criteria for proposals and incorporates sections for describing the research problems being investigated and for justifying the use of supercomputer systems for the computations involved.

As proposals are received by ER programs, they are evaluated and ranked by the ER program managers. National laboratory contractor proposals are folded in with all others and evaluated on a consistent basis.

Within each ER program area, there is a single supercomputer access program representative who compiles these proposals for the program, e.g., Basic Energy Sciences, and who represents the program on a cross program

allocation committee. This allocation committee then works with these compiled proposal summaries which have been evaluated within each program with regard to scientific merit, priority for meeting overall program goals and individual supercomputer justification. Finally, this allocation committee determines resource allocations based on these rankings and on the total resources that will be available for the given fiscal year.

Question 6: What portion of the Office of Energy Research's funding is designated for training and support versus hardware and communications equipment?

Answer: The Office of Energy Research funding for Supercomputer access for FY 1985 is $30,100,000, of which $8,065,000 or 26.8% is designated as training and user support. This funding is broken out as follows:

1. The National Magnetic Fusion Energy Computer Center funding for FY 1985 is $23,100,000, of which $3,100,000, or 13.4% the Department designates as training and user support.
2. The Supercomputer Computations Research Institute funding (at the Florida State University) from the Office of Energy Research for FY 1985 is $7,000,000, of which $4,965,000, or 70.9% FSU designates as training and user support.

Question 7: What are your nerar- to mid-term plans with regard to supercomputer procurement? Do you have any plans to establish additional supercomputer center?

Answer: During FY 1984, the Office of Energy Research (ER) initiated plans to expand the existing Fusion Energy supercomputer access plans to the non-fusion research community. At that time, a thorough analysis indicated that utilizing the National Magnetic Fusion Energy Computer Network and implemen ting the departmental plans with a Cray X-MP/22 computer system constituted the most cost-effective and expeditious method of providing supercomputer access to the large geographically dispersed non-fusion ER research community. The Cray-XMp/22 system acquired during FY 1985 on an interim basis, now needs to be replaced by a more powerful Class VII system. When funding is available, DOE intends to procure such a system competitively and replace the currently used Cray X-MP/22 at the National Magnetic Fusion Energy Computation Center. This Class VII system would again be funded through the Basic Energy, Applied Mathematics Program for use by the HENP, BES, HER, research communities.

At this point in time, ER is continuing to evaluate the needs of its research community and has not as yet determined whether continued use of the NMFECC or whether establishing and additional center would be the most advantageous method of supporting this requirement in the future. The benefits in support of the ER mission, cost-effectiveness of providing access to supercomputer systems of differing architectures, and the case of cross program collaboration are now being weighed to determine the most advantageous approach.

For the Magnetic Fusion Energy Program in 1985, a Cray 2 computer system was installed. Currently, we believe that this Cray 2 system will be adequate to meet the supercomputer computational needs of the fusion community in the near-term. We have not determined what the additional computer needs of the fusion community will be for the long-term.

Question 8: Europe has a number of supercomputers, and Japan has recently purchased a Cray. Do you have any plans for establishing international networks so that researchers around the world could use our supercomputers, and vice versa? What are the problems or difficulties with establishing such a network?

Answer: The Office of Energy Research (ER) supercomputer access program provides supercomputer services to researchers who are funded through a contract or grant with ER. At present, the number of researchers who are on contract to ER and who receive supercomputer services is approximately 4,000. These researchers are geographically dispersed throughout the United States and currently access the ER supercomputer systems via the National Magnetic Fusion Computer Network (MFENET). ER currently has two international agreements for joint research projects which involve a small amount of sharing of supercomputer resources with Japan and with Germany. Communications links to computer installations to these countries are now under development. Full network integration is, however, not anticipated in either case since gateways between the MFENET and these other networks appear to be adequate and to offer better internal controls. ER is also pursuing, through the FCCSET committee possible connections of the MFENET to networks supported by other agencies.

Experts in data communications agree that the use of multiple networks interconnected through highly functional communications links is the current trend throughout the telecommunications industry and the most viable approach for meeting future data communications requirements. (This trend is evidenced by the recent partitioning of the ARPANET into two separate networks, MILNET and ARPANET). Hence, current plans for ER data communications are to continue to operate the existing MFENET to meet the ER data communications requirement, to evolve this network toward more general functionality, new technologies, and emerging network standards, and to interconnect to other agency and international networks through highly functional communications links.

NATIONAL SCIENCE BOARD
WASHINGTON, D.C. 20550

August 5, 1985

Honorable Don Fuqua
Chairman
Committee on Science and Technology
Washington, D.C. 20515

Dear Mr. Chairman:

Attached are the answers to the questions which were forwarded with your letter of June 25. I hope they will be useful to the Subcommittees on Energy Development and Applications and Science, Research and Technology.

Sincerely,

Mary L. Good

Mary L. Good

Attachment

Response to the questions posed by the House Science and Technology Committee in the letter from Rep. Fuqua to Dr. Good on June 25, 1985.

1. Q: What coordination is taking place among the DoE, NSF and DARPA supercomputer programs?

A: Through the Federal Coordinating Committee for Science, Engineering and Technology (FCCSET), representatives of these three, and other, agencies regularly meet and coordinate their respective supercomputer plans. Several joint projects are in the planning stages, such as the joint NSF/DARPA networking plans, the summer training institutes, and the plan to exchange supercomputer resources between the supercomputer centers funded by NSF, DoE and NASA.

Q: What other agencies are involved in this area?

A: Besides the NSF, DoE and DARPA, several other agencies have substantial interest in supercomputers and their applications. These include NASA, Commerce (through the supercomputer centers at NBS and NOAA), and HHS (through the anticipated supercomputer center funded by NIH).

2. Q: How are NSF priorities and allocation of funds for basic computer science research (architecture, software, etc.) and the Advanced Scientific Computing program established?

A: Through various studies, such as the Lax and Bardon-Curtis reports, and through the advice of advisory committees, NSF determines what are the needs of those parts of science and engineering which have important contributions to make, and funding allocations are made accordingly. Both computer research and advanced scientific computing have received large increases recently since they are both important for the advancement of basic knowledge and national needs.

3. Q: On page 3 of your testimony, you refer to the supercomputer that will be transferred from NASA along with your discussion of interim centers. (a) Do you expect the university which receives the NASA computer to establish a National or interim center?

A: Current budget levels do not allow for providing operating and maintenance funds to the Pittsburgh center, which will receive the NASA supercomputer. These funds will be covered from local sources, and the local management will have control over the allocation of the resources of this center. Initially, therefore, the Pittsburgh center will be neither an interim nor a national center.

Q: (b) Does NSF plan to upgrade the NASA computer to be comparable to the National Centers' facilities?

A: If future budgets allow, we will consider an upgrade to a state of the art supercomputer for the Pittsburgh center, and the initiation of a National Facility there to serve the entire science and engineering research community. However, at the current levels, this will not be possible.

4. Q: (a) How will the allocation of time at NSF supercomputer centers be addressed to avoid local users dominating the time at a specific center?

A: All grants of supercomputer time, including those to local users, must go through a peer review procedure, which will insure that local users will not dominate the time at a specific center.

Q: (b) What policy or criteria will be established to insure that national goals are achieved?

A: Two of the most important national goals are the advancement of scientific knowledge, and the training of students and young researchers in the use of supercomputers. Therefore, these will be the most important review criteria under which the excellence of proposals are judged.

Q: (c) How can all disciplines have fair access?

A: Any review procedure will be carefully balanced so that disciplinary biases are minimized.

Q: (d) How, and by whom, will interdisciplinary priorities be set?

A: The interdisciplinary priorities will be set by the Office of Advanced Scientific Computing on the advice of, and subject to the review by, the Program Advisory Committee. This committee, which meets twice a year, consists of fifteen experts from all branches of science and engineering.

5. Q: The initiation of the NSF National Centers represents a long range commitment over at least five years. (a) What is NSF's average support level for each center in FY1985?

A: In FY1985, the average support for the first four centers (San Diego, Illinois, Cornell and Princeton) is $5.7M.

Q: (b) What are the funding level projections for the centers in FY 1986 and beyond

A: In FY1986, the average support for the four centers has been set at $7.33M, contingent on satisfactory progress and the availability of funds. The total amount which has been committed for the next five years is $183.1M. This amount includes only three years' commitment to the Cornell center, which is more experimental in nature than the other three. If the Cornell center is continued into the fourth and fifth years, the total support will be slightly more than $200M.

Q: (c) Will years three and four of the Centers' operation require greater support from NSF than the first years?

A: Yes, the average level per center in FY1987 will be $[illegible]M, which is $1.7M per center higher than in FY1986, and $[illegible]M higher than in FY1985. We expect the centers to start to provide service midway in FY1986, so that FY1987 will be the first year of full operation. It is therefore natural that the third year should be higher than the first two. The fourth and subsequent years will be approximately the same as the third, with allowances for inflation.

Q: (d) What proportion of the NSF Advanced Scientific Computing funds will be expended for Center operation versus establishing a national network and local work stations?

A: In FY1985, 78% of these funds will be expended on Center operation, versus 17% on networks and local access equipment. At the FY1986 request level, these proportions are 85% and 11% respectively. The difference is due to placing a higher priority on funding the full operation of the Centers than on continuing to fund the local access equipment. This cost will have to be absorbed by the individual programs in the research directorates. Future funding proportions will depend on total budget levels and priorities.

6. Q: (a) When do you anticipate a NSF national network will be operational?

A: We expect the Phase 1 network (which will be a combination of existing networks like ARPANET and BITNET, consortium networks, and direct lines between centers) will be operational by early 1986. The Phase 2 network, which will be newly-designed, will be operational within three years. Several pilot projects and development grants to test new technology will be necessary before the Phase 2 network will become a reality.

Q. (c) What resources will be needed to enhance networking in the short term?

A. Adequate funding, cooperation from other agencies and networks, and sufficient staff [illegible] to manage the networking program, are the most critical resources needed to make this program a success. So far, the FY1986 funding, and FY1987 requested level, are sufficient to start the Phase I network, and to begin a small effort on developing Phase II supercomputer access. However, there will need to be a significant increase in FY1988 to meet the goals of the national network.

Q: (d) What resources will be needed in the long term to establish a national network including workstation access and local networks within universities?

A. By 19[illegible], we will require a minimum of $10M in the networking program in order to provide access for the science and engineering academic community to the NSF supercomputer centers. If we were to include other networking needs and to make a truly national network which would include the industrial and governmental sectors, then the requirements would be more than double this figure.

Q. (e) What will be NSF's role in establishing national and local networks and how will these needs be shared and coordinated with other Federal agencies and the private sector?

A: NSF has already assumed a lead role in coordinating the network activities in universities. This involves funding connections between local, regional and national networks, and encouraging compatibility between them. Whenever possible, the needs of other agencies will be taken into account, and joint projects [illegible] will be set up [illegible] private sector in several future projects.

7. Q. (a) Are NSF grants for supercomputer time limited to NSF supercomputer sites (interim or national centers)?

A: At the moment, grants for supercomputer time are made at the six Phase I centers. The national centers are not yet operational. Also, both individual research programs and the Office of Advanced Scientific Computing do provide some support for access to other supercomputer centers. This is for researchers who have particular special needs. Furthermore, through the FCCSET supercomputer committee, there is a plan for the agencies to exchange supercomputer resources on a quid pro quo basis.

Q: (b) If so, what impact will this policy have on the formation of new supercomputer centers funded by states, universities, and industry?

A: The NSF policy should have a positive effect on the formation of new supercomputer centers, since researchers will be able to apply for funding to access such centers.

Q: (c) What policy options could be considered to increase flexibility and still use the NSF centers to the maximum extent?

A: One possible option would be that described in the answer [illegible](d). Resources could be exchanged among many different supercomputer centers without the exchange of funds, subject of course to the approval of the peer review process.

3. Q: Experience at the interim sites has shown that the training and ease of access significantly affect the rate at which new users actually use their allocation of supercomputer time. What proportion of the NSF centers funded have been designated for training and support versus hardware and communication?

A: All of the interim centers have established training programs as part of their agreement with NSF. This procedure will be continued for the National Centers. We have also funded three summer training institutes which will be held during August of this year at Boeing Computer Services, the University of Minnesota and the National Center for Atmospheric Research.

4. Q: (a) Does NSF plan to treat and manage the five National Centers as a single resource or a collection of independent centers?

A: Since each of the centers initially will be quite different in configuration and scope, we will treat them as independent entities as much as possible. As with other NSF facilities, local management (with close monitoring by NSF) is the most efficient means of allocating resources.

Q: (b) What are the advantages of the single resource versus independent resource concept?

A: The single resource concept will be an attractive alternative when the national network has been established and compatible operating systems and other software have been installed. It will then be easy for a user to transfer between centers, and the incompatibility between different types of supercomputers will cease to be a major factor.

[illegible] of the National Centers are [illegible] resources, can an NSF [illegible] function effectively [illegible] project manager, or will a contract management consortium, similar to ocean drilling and astronomy center contracts, [illegible]

As two of the new national [illegible] centers will be run by consortia. These are the San Diego Supercomputer Center, which is allocated to a consortium of [illegible] universities and research institutions, and the John Von Neumann Center in Princeton, NJ, which is managed by a consortium of [illegible] universities. The connections between the centers and their consortium members are part of the NSF networking plan. Eventually, as we progress toward the second phase national network, we will arrange for a network management contractor to take over the detailed running of that part of the program.

DEFENSE ADVANCED RESEARCH PROJECTS AGENCY
1400 WILSON BOULEVARD
ARLINGTON VIRGINIA 22209

JUL 17 1985

Honorable Don Fuqua
Chairman, Committee on Science
and Technology
U.S. House of Representatives
Suite 2321 Rayburn House Office Building
Washington, D.C. 20515

Dear Mr. Chairman:

Enclosed are the answers to the questions asked by the Members of the Subcommittees on Energy Development and Applications and Science, Research and Technology which resulted from the hearing in Tallahassee, Florida on 10 June 1985.

I would like to thank you and your Subcommittees for the interest given the DARPA Strategic Computing effort. This effort is a coordinated attempt by various research groups to provide this nation with dramatic advances in microelectronics, computer architecture and artificial intelligence. This effort will offer not only the military but the commercial world opportunities to use the computer in more significant roles than ever before considered.

If you have any questions regarding these answers or the program in general, please contact Mr. James C. Goodwyn (202) 694-1440.

Charles Buffalano
Charles Buffalano
Acting Director

Enclosure: a/s

RECEIVED
JUL 19 1985
COMMITTEE ON SCIENCE
AND TECHNOLOGY

QUESTIONS AND ANSWERS

1. What coordination is taking place among the DOE, NSF and DARPA supercomputer programs? What other federal agencies are involved in this area?

DOE, NSF, DARPA and the other agencies involved in supercomputer research (principally NASA, CIA, NBS and other parts of DoD) normally coordinate their research programs on a regular basis at the program manager level. In the last two years, an additional level of coordination was established by OSTP as part of the Federal Coordinating Council on Science, Engineering and Technology (FCCSET). In particular three panels were established under FCCSET: (1) on supercomputer procurement, (2) on supercomputer access and (3) on computer research. The first two panels were subsequently merged into one. All the relevant federal agencies are directly involved in both FCCSET panels.

2. What are the critical areas of supercomputer R&D?

A few examples of critical supercomputer R&D for DARPA are:

A. Achieving very high speed computing for machine intelligence technology and its application.

B. Developing architectures designed to support effectively numeric and symbolic computing.

C. Developing languages and software systems suitable for expressing and and implementing parallelism.

D. Developing techniques for high speed random access secondary memory systems.

E. Determining how to utilize very high speed computing for future defense systems.

3. What is the private sector doing in supercomputer R&D, and how are the federal programs coupled into these private sector activities?

The private sector has had only a limited investment in supercomputer R&D, with expertise residing principally at three U.S. Companies (Cray Research, ETA Systems and Denelcor). Recently, other companies such as Amdahl, various Japanese Companies and a set of mini-supercomputer manufacturers have begun to sell high and low end machines on the market. These are basically conventional uni-processor architectures or very small scale multi-processors. The federal programs are assisting by buying their machines to offset development costs and by learning how to use multi-processors for specific applications as in energy research or aeronautics applications.

4. We hear of the possibility of having a "supercomputer on a chip" that may have the computing power of today's large machines at a fraction of the cost. Is this a "pie in the sky" notion, or is it a real possibility, and if it is real, how soon might we see such devices.

A supercomputer will undoubtedly be implemented as a wafer-scale system by the end of the century, but the entire package will probably be significantly larger than today's microcomputer chips. To be useful, the system will require very large amounts of memory, other peripheral equipment and may require cooling apparatus. However, the overall system could be significantly lower in cost than today's largest machines for the same equivalent computing power and occupy only a fraction of the space.

5. (a) Could you describe for us DARPA's Strategic Computing Program?
   (b) What are the Program's goals, and what is the Program's funding level?

The attached plan and first annual report describes the DARPA Strategic Computing Program. Briefly, the program is developing a new class of machine intelligence technology which is intended for applications to critical defense problems. However, the technology itself will be generic, will be produced by industry or joint university/industrial efforts and industry is expected to make it widely available.

6. (a) What level of funding for supercomputing R&D is DARPA providing to universities? (b) How much to industry? (c) Is DARPA funding any research at government laboratories?

The DARPA Strategic Computing effort is a coordinated research effort between industry, universities and government laboratories. The FY 84-85 funding proportions of each group relative to the total Strategic Computing budget are as follows:

Universities	--	$ 45.0	37%
Industry	--	51.0	42%
Government Laboratories and Research Centers	--	10.0	8%
Non-Profit Institutions	--	2.5	2%
Other	--	13.5	11%
		$122.0	100%

PANEL 2: SUPERCOMPUTER CENTER OFFICIALS

Mr. FUQUA. Our next panel will be Dr. Larry Smarr, professor of astronomy at the University of Illinois in Urbana; followed by Dr. Ken Wilson, professor of physics at the Department of Nuclear Studies at Cornell University in Ithaca; Dr. John Killeen, Director of Magnetic Fusion Energy Computer Center at Lawrence Livermore National Laboratory in California; and Dr. Robert Johnson, dean of graduate studies and research here at Florida State University.

Dr. Smarr, we would be pleased for you to begin.

[The biographical sketch of Dr. Smarr follows:]

LARRY L. SMARR

Larry Smarr, 36, is professor of physics and of astronomy at the University of Illinois at Urbana-Champaign and director of the university's National Center for Supercomputing Applications. An internationally recognized astrophysicist, Smarr was the lead scientist in a proposal to the National Science Foundation to establish the center at Illinois.

A U. of I. faculty member for six years, he is one of the pioneers in the movement to obtain federal support for supercomputing power at U.S. universities for basic scientific research. His 1982 paper on "The Supercomputer Famine at American Universities" is regarded as a landmark in that movement. It was written after he had to travel to West Germany to gain access to an American-made supercomputer for research, and it resulted in his successful proposal to the NSF. He was one of the major co-authors of the "Press Report," a seminal national study of future U.S. computational needs in theoretical physics.

Smarr earned bachelor's and master's degrees from the University of Missouri, a master's at Stanford University, and a doctorate from the University of Texas at Austin. He taught and conducted research at Stanford and Texas, and at Princeton and Yale universities. The three years before he joined the U. of I. faculty in 1979, Smarr was a Junior Fellow in the Harvard University Society of Fellows.

A consultant to several government and professional scientific agencies, Smarr was co-founder of the Illinois Alliance to Prevent Nuclear War, and continues to be active in the debate on nuclear weapons and arms control.

Esquire magazine named Smarr one of the "Best of the New Generation, Men and Women Under Forty Who Are Changing America." His views on supercomputers and science have been quoted widely, including in the *New York Times, Wall Street Journal, Time, Business Week, Los Angeles Times, Chicago Tribune, Science* and *Science News.*

STATEMENT OF DR. LARRY L. SMARR, PROFESSOR, ASTRONOMY DEPARTMENT; AND DIRECTOR, NATIONAL CENTER FOR SUPERCOMPUTING APPLICATIONS, UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN, CHAMPAIGN, IL

Dr. SMARR. Thank you very much. It really does give me a great deal of pleasure to be testifying to the committee, because I worked for a number of years, back about 5 years ago, as was mentioned earlier in the Press Report, and I think the really striking initiative your committee took last year in going beyond the President's recommendation and in fact initiating out of subcommittee, a $100 million program which I think is the correct level, showed a real vision and also ability on Congress' part to listen to scientists in this country and to hear our cry that things were not the way they should be.

The National Science Foundation has also of course responded and is showing a great deal of leadership now in rectifying this problem.

I have provided for this committee, a reprint of a recent article from *Science* magazine that I wrote, for several reasons, one, to remind you that the push behind this effort is coming from science, but I want to give you not from the magazine but from the basic research needs of our country and I want you to have an image in your mind.

Too much of this testimony is on paper and that is not where the revolution is at. This sort of visualization of a complex phenomenon is why the supercomputers are needed coupled with work stations. But another important fact is this calculation was done in units on the Cray-1 supercomputer.

The good news is even though this has just come out in *Science*, this week the Scientific Advisory Committee to the NSF will see a movie made at Digital Productions, one of the new phase 1 centers sponsored by your initiative and the NSF, and that the calculations of this problem are now being done in this country and in the fall will be transferred permanently to the national centers.

So I think we have seen an incredibly rapid turnaround here and the scientific community owes the Congress a great deal of gratitude for its moving so rapidly.

I would like to ask that the prepared testimony I have be entered into the record.

Mr. FUQUA. Without objection, we will make it a part of the record.

Dr. SMARR. Thank you very much.

I will try to hit on a few of the key concerns that I think all the national centers have and I will illustrate as I go along with a couple of examples from our particular center.

The supercomputer I think Dr. Good mentioned, we have not quite got ours yet—it will be August—but we will be up and running quickly. I think something that is very important to keep in mind is that the next 5 years will be radically different in the supercomputing technology than the previous 5 years.

The Cray 1 supercomputer was introduced in 1976–77 and until last year was the reigning supercomputer along with the CDC–205, and therefore, over the last essentially 8 years, no increase in speed at the high end has occurred.

But now we are entering the age of multiprocessor supercomputers where one can put 2 of those, 4, 8, 16, as far as we know without limit, into the thousands, processors inside the same box.

And so in the next 5 years, we are going to see a virtual exponential increase in the power of these machines in which by the end of the 5-year program our machine will be 50 to 100 times as powerful as the Cray 1 and continuing to increase every 2 years after that.

So we are in a brandnew ball game and the new initiative could not come at a more crucial moment in the history of our technology.

The approach at our center, and I think at some of the others is to be at once conservative in that we are trying to build on the—stand on the shoulders of giants, as it were, the thousands of man-years that have gone into developing the software and the hardware in these supercomputer in the national laboratories, DOE laboratories, particularly in our case with the Cray supercomputer,

and we want to put that—transfer directly, conservatively, reliably into the hands of the scientists who need it.

We are also being quite radical in that we want those supercomputers to be much more useful than they are today. Supercomputers in isolation are not that useful. What we have to look at is the entire computer revolution as a single organism in which the personal computer, the network and the supercomputer is one entity and that entity is what is used by the scientific community and by industry to solve problems and develop productivity, and that is what we are going to try to establish.

One of the real problems is that this technology is going to be much more difficult for some scientists to learn how to use on their own than the previous versions of supercomputers.

And that is because previously with the vector supercomputers, essentially a scientist on his own could learn how to optimize his computer programs to run most efficiently on say a Cray or a Cyber, and now because it is much more complicated, because of the multiprocessors, he has to define how you are going to divide the program up among these various processors to run simultaneously.

So I think you are going to see the emergence of a new prototype scientific research group in which the computer professional is much more a member of the team than has previously been the case.

In Illinois, we have both the center—National Center for Supercomputing Applications, which I am director of, funded by the NSF and the State, as well as a different supercomputer center, one that Dave Kuck heads, the Center for Supercomputer Research and Development, which is primarily DOE and some NSF support as well.

And that center really is a computer science think tank, to look at these new kinds of architecture that are developed and these new kinds of problems and figure out how can you take previous work and restructure it. I think that has got to be a major part of the science that we are doing and so the science has got to be expanded and these national centers in the universities are apparently crucial for that expanded activity to materialize—new human bonds and working relationships must be developed.

Now there is a lot of talk about networks. In the testimony, I use the analogy of the existing networks of, say, water distribution. Previously I had used the Interstate Highway System and I am glad to see Senator Gore referring to that. You are talking about a very complex entity when you talk about a national network.

When you think of large reservoirs outside of cities and the great pipes taking it down to different cities all over the State, it gets into the cities and breaks up into neighborhoods and then in the neighborhoods they break up into the houses and even in the houses, they break up and finally you get to the faucet which you can turn on and get a drink of water out of instead of being drowned if they open the spillway on the reservoir.

I think that is exactly the problem we are facing—it is no less complicated and it will be no less an effort to build this network. Fortunately, supercomputers are not the only game in town. The personal computer revolution is, I think, absolutely crucial to the

effective use of supercomputers because it is the personal computer that will become the faucet.

The reason for buying personal computers in the homes—and there are some corporations with 10,000 personal computers currently in place—is because that is a direct faucet on the information distribution system to get an amount of information that a person can digest.

The question is how do we go from the supercomputers which are the great reservoirs, the generators of this information, to the personal computer, and that is the networks' function.

So I think it is an enormous undertaking. I think what we are seeing in the centers and what we are doing in Illinois is we are currently not only working on how to make the faucets work better, but also we are developing a whole system of diameters of pipes for carrying information; everything from the 1,200 bits per second which you can get on a telephone line up through the ETHERNET-type environment of, say, 10 million bits a second.

In fact, we are cooperating with a recent initiative at Los Alamos to go all the way to gigabaud, a billion bits per second on fiberoptics. That is a factor of a million in transmission speed.

All of those pipes are going to be needed to build this great network. But that is an experimental developmental effort and you go from those faucets and that huge pile of pipes that is developed here and in the industry and in the other centers, to putting it together to build a national network, which is an awesome undertaking.

I think that we have to decide what is the NSF's correct role in all that and what is the private sector's, what is DOE's. We are only a small fraction of the information to be carried on the information distribution system. I do not have an answer for you, but I think it is important to realize how awesome the undertaking is and yet how we cannot avoid it. The future of this country and the world depends on that happening and it will happen.

The other thing I am concerned with as the director of a computer center, is that the scientific community continue to support this effort and that will only happen if every researcher in this country who has an idea and it goes through peer review and is accepted, has access to these facilities.

Now we are not funded in our center, for instance, by the NSF, to provide that with telephone line rates and that sort of thing. Something has got to be done to keep the community behind this program and make sure that they get access to it, and that is the network, and that is why that is the next important goal.

Let me turn to education. One of the reasons I think that we all want these centers in the universities is because of the educational possibilities. At Illinois, for instance, we have 400 graduate students in computer science, 300 in physics, 1,000 in engineering; all those students will be able to get small amounts of time on the supercomputer.

I have a high school student who is a real hotshot who is working this summer with me on doing this sort of color graphics on the supercomputer.

In fact, our university high school has requested a port for the supercomputer in the high school and out of my discretionary fund

they will get it. So we are very committed to education at all levels. But I think the fact is we probably need more education than the students do.

The students will come in and pick up immediately. And so we are building an academic program around the supercomputer center at Illinois and in fact a new kind of scientific institute in which you can bring scientists, engineers, scholars, everyone and get them excited and working in a work station environment and then take that and transfer it back into their fields.

I think that is very important to realize. If you look at the bottomline numbers that your committee and the NSF is putting into these centers, ours is about 60 percent funded by the NSF, the rest is State, university, and vendor funded. That is a small fraction of the actual program because it is the entire academic program that is being funded by the university that does not appear on the books anywhere that is really developing new courses, et cetera.

I think that the industry must come into the mix now. There is a very good partnership developing now for Federal, State, university, and computer manufacturers working together to make these centers. But the industrial users of these machines are not yet in that mix and I would like to see them come in as a full share by expanding the centers to adequate capacity to deal on the applied research topics that the industry is interested in, but working together, bringing their people from industry and living in universities in these centers and becoming an overall part of this research effort, not just the basic research and applied research, but make it all happen together.

I think, just in summary, the things that I see that are most important is that the centers that are now started must be sustained at the level of state-of-the-art, and that means that the out-year budgets, the 1987 to 1990 fiscal years, I think you have to look very carefully to make sure that there are adequate funds to maintain these centers, because we have pulled down our budgets in these first years to work with the Congress because of the deficit problems.

That means it is more significant perhaps than it would have been otherwise in getting these centers to where we say we are going to be. The national network is really the next challenge and that is something that we have got to make sure there is adequate funding and we have got to make sure what the correct role is that we are funding the NSF to do in the overall building of the national network.

Finally, I think this question of capacity, I think there has got to be at least three levels of attack. The supercomputer should be used like the Fermilab or the national observatories to attack those problems that cannot be done otherwise, really push the frontiers of knowledge; but we are going to find very quickly that research groups need a 1,000 hours a year and there are only 7,000 hours a year and already there are researchers using a 1,000 hours a year of supercomputer time. They obviously are not going to be able to get on these machines, and they need dedicated minisupercomputers or something, but they need their own dedicated equipment.

And then finally, we need to faucets, every researcher in this country needs to be working at least on personal computers, if not

on the new scientific work stations which will ultimately be replacing the minicomputers of the 1970's. I think those programs must be balanced with the network so that when we develop the new coputational facility, it really will make the breakthroughs in science that we all want to see happening.

Thank you.

[The prepared statement of Dr. Smarr follows:]

University of Illinois
at Urbana-Champaign

National Center for Supercomputing Applications
153 Water Resources Building
605 East Springfield Avenue
Champaign
Illinois 61820

Dr. Larry L. Smarr
Director
217 244-0072

DR. LARRY L. SMARR

Director, National Center for Supercomputing Applications

University of Illinois at Urbana-Champaign

CONGRESSIONAL TESTIMONY

Delivered to a joint field hearing of the

Subcommittee on Science, Research and Technology

and the

Subcommittee on Energy Development and Applications

on Federal Supercomputer Programs and Policies

June 10, 1985

Tallahassee, Florida

Because of the bold initiative taken one year ago by the Committee on Science and Technology and the National Science Foundation, I am able to address you today as the Director of the recently established National Center for Supercomputing Applications (NCSA) at the University of Illinois at Urbana-Champaign. This new center, one of the largest projects in the history of the University of Illinois, is jointly funded by the Office of Advanced Scientific Computing at the National Science Foundation, by the University of Illinois, the State of Illinois, and CRAY Research, Inc.

I would like to share with you some of the hopes and dreams we have for what this center will accomplish, not only for basic scientific research but the future of industrial America.

We have designed this Center in a manner which is at once both very conservative and quite radical. This is because we want our Center to be able to deliver reliable service as soon as possible, while at the same time we intend to be very innovative in developing new ways to make supercomputers much more accessible than they are today.

We are basing the major components of the facility on the existing strength of the hardware and the software which has been developed with thousands of man-years of effort at the national DOE laboratories.

In particular, we are using the CRAY Research, Inc. line of supercomputers. This will allow us to run the wide universe of software for the CRAY which has been developed over the last ten years. This is a very conservative approach, because without a widely available variety of

debugged and reliable software a fast machine is really of very little use to a wide audience. The operating system will be the CRAY Time-Sharing System (CTSS), which was developed at the national labs, and which allows a number of users to simultaneously work on the CRAY supercomputer. There are a wide variety of software productivity tools available with this operating system. Finally, we will use the Common File System (CFS) that Los Alamos National Laboratory has developed as the software which keeps track of the user files. This will enable us to keep track of the awesome amount of information that the CRAY generates. Using CFS, a user, within seconds, can select a file from the several terabits (one thousand billion bits) of information available on the mass storage unit.

Thus, the most crucial components of the central facility: the supercomputer, its operating system, and its file system, are proven reliable technology that without question can deliver the goods from day one.

Although our choice of supercomputer is very conservative, the University of Illinois has pioneered the radical concept of rapid upgrades in the power of that supercomputer.

This is possible because we have entered the age of the "multiprocessor" supercomputer. This is a supercomputer, composed of several identical computers, which "gang tackles" a problem. This concept evolved from the innovative ILLIAC IV multiprocessor supercomputer designed at the University of Illinois in the 1960's. However, whereas all processors of the ILLIAC IV had to work in unison, today's multiprocessors

have each processor independently work on a separate part of the problem. Illinois will start with the CRAY X-MP multiprocessor, which contains two supercomputers inside it, each more powerful than the previous CRAY-1 single processor supercomputer. We will upgrade our supercomputer every one to two years. Thus, sequence of machines will have to four, eight, and finally sixteen processors at the end of the five year NSF agreement. Because each individual processor also gets faster in each new generation, in only five years we will have moved to a machine that is 50-100 times faster than the current Cray-1 supercomputer.

All future fast machines must have multiprocessor architecture if they are to be competitive. Unfortunately, almost all computer programs are written for single processor computers. Therefore it is crucial that our basic research community learn immediately how to use multiprocessors. At Illinois we have established a second supercomputer think tank, the Center for Supercomputing Research and Development, whose mission is to explore the computer science of these new problems. That Center will be directed by Professor David Kuck, one of the pioneers of supercomputer architecture and software. The Center will build experimental multiprocessors, develop new software and operating systems for multiprocessors, and provide tools to restructure old FORTRAN computer programs to run on these new machines. The NCSA will work closely with the CSRD at Illinois to help transfer these new ideas to the national basic research community.

Thus, the University of Illinois is unique in having two supercomputer centers. One, the National Center for Supercomputing Applications, will take the best hardware and software from the marketplace and the government labs and make it available to the national basic research community. The second, the Center for Supercomputing Research and Development, will experimentally build prototypes of the supercomputers and their software which will become commercially available in the 1990's and beyond.

The good news is that enormous computational power is becoming available to researchers in the near future. However, the bad news is that the great supercomputers are very unbalanced with respect to delivering information to a human being.

Perhaps a good analogy is that of an enormous water reservoir near a large city. When the spillway is opened, an enormous volume of water is released. This must be channeled through a series of large distribution pipes to various parts of the city, then smaller pipes deliver it to neighborhoods, within the neighborhoods smaller pipes deliver it to houses, and finally in a house an individual can turn on a faucet and get one or two drops of water at a time. The flow of information from a supercomputer is very similar. Unfortunately, in our country today we have very little other than the large spillway available. At the University of Illinois we are trying to help develop a distribution system for this information so that a single user can get a "few drops" of information at a time.

The key to success in this area has two parts. First, is the personal computer (PC) revolution. This rather unexpected development of technology has perfected the "faucet" of our delivery system. The PC is the device that all of us, whether we are at home or at work or in the laboratory, have had to learn to use to retrieve information. As such, researchers have become familiar with the "user-friendly" interface on their PC. At Illinois we are undertaking research to develop software so that you can run the CRAY from your personal computer in such a way that you think the CRAY is "inside your PC." Our ideal is that within a few years a researcher will be able plug in his or her lap top PC into a telephone anywhere in the country and run the CRAY at the University of Illinois. This will make access possible to supercomputers on a vastly expanded basis.

The second major part of the distribution system that must be developed is that of the network. We at Illinois are striving to develop a large range of data transmission speeds, in analogy to the large range of diameters of pipe for the water distribution system, so that the appropriate technology can be used at the appropriate place in the distribution system. As I mentioned before, we are trying to make the CRAY very useful at extremely low transmission speeds such as 1200 bits per second of the normal PC modem. We will also be hooking research groups, in which each scientist has a powerful scientific workstation on his desk, together at ten million bits per second, one thousand times faster than a modem. Finally, we will be experimenting with fiber optic links, which promises to be able to go yet a thousand times faster, up to billions of bits per second.

Thus, the University of Illinois will be experimenting with "information flow pipes" whose capacities differ by a factor of over one million.

For each of these transmission speeds there is an appropriate personal computer or scientific workstation which can handle that information flow and perform tasks from simple text processing at the low end to elaborate three dimensional color simulation movies at the high end.

This coupling of the personal computer revolution to the supercomputer revolution is of utmost importance if we are going to make the most productivitive use of these scarce resources.

The Congress must understand the enormity of the task of creating a "seemless web" National Network, hooking all researchers together at high speed, starting with the new ideas and technologies which we, other universities, national labs, and private companies are developing.
The only part of this network that our national center is funded for besides the development and experiments on networking is to set up a "front end" computer at Illinois to use existing telephone lines and to become part of DOD's ARPANET.

We look to the Congress to appropriate enough money to make the National Science Foundation's plans for a basic research national network a reality. However, the effort will be a major one lasting many years, just as have been the previous networking efforts, which joined this country together with roads, telephones, television, and electrical power lines.

The wisdom of Congress and the NSF was to place these National Centers in some of the great research universities of this country now, so that we can get going, while we wait for the national network to develop. This will allow for the training of thousands of new students on these new technologies. These students will go to other universities, to industry, and to government laboratories carrying skills and new approaches which will hasten the transition to supercomputing in their new jobs.

But at Illinois we believe that merely delivering supercomputer cycles and training students is not an adequate response to the supercomputer challenge. We feel that today's methods of using supercomputers in research are not nearly productive enough. We must couple all parts of the computer revolution together if we are going to increase our scientists ability to get the most insight out of each supercomputer project.

Therefore, as part of Illinois's cost-sharing we have established an Intellectual Center at the hub of the NCSA. It is a building in the heart of the campus which will bring together from the nation scientists, engineers, social scientists, scholars, and computer professionals in a new

type of scientific institute. This institute will be broadly multidisciplinary, with a mission to seek out those common problems researchers experience in trying to use their common computational tools.

These researchers will work in a modern comprehensive computational environment, unprecedented in a scientific institute. Each desk will have a personal computer or a scientific workstation networked to the supercomputer. Laser printers will produce typeset quality output. Users will be able to view color movies of their simulations directly on their workstation. Electronic mail will link all researchers in a common dialogue. The goal is to experimentally determine how much scientific productivity can be improved by removing the technological bottlenecks which have always before hindered our rate of progress.

By having researchers working side by side with nationally selected computer professionals in this state-of-the-art computational environment, we will be able to find innovative solutions to many of these problems. In addition, we will have the "critical mass" of investigators necessary to "dig deep" in a number of frontier research areas in science and engineering. The institute will have a strong national visitor program so that these new tools can be widely disseminated among the research community.

The best way to develop those new technologies which make supercomputers more useful to researchers is to have the researchers themselves intimately involved in the design of the technologies.

Illinois is committed to the rapid transfer of the new technologies we develop into the private sector. It is my belief that the NSF centers can have a major impact on the competitiveness of American industry.

The future for our country lies in the information age and high technology. Although most corporations know this, they are not in the business of performing the type of experimental basic research that the centers can. Therefore, I believe that the centers may become the nucleus for a new type of partnership between Government, Universities, and Industry to assure that these technologies can be developed and applied to American industries in time to allow them to remain competitive in the global marketplace.

At Illinois we pursuing discussions with major research corporations to see if a major enhancement of the NCSA would be possible. We do not believe that these corporate partners will replace the NSF dollars for university basic research. Rather, we see them as augmenting the existing program in ways that help transfer the new technologies directly into industry. The enhancements in scientific productivity that we forsee making in the NSF centers should translate into engineering and manufacturing productivities in the marketplace. This transfer will not happen without a new framework for transferring the technology. I view the university centers as the correct "neutral ground" for this transfer to occur by developing human partnerships between research scientists in academe and in industry working together in the expanded centers.

If the centers are successful in obtaining new industrial funds, the NSF dollars Congress has invested will be enormously levered by the increase in national industrial productivity.

The University of Illinois is undertaking one of the largest projects in its history in building this national supercomputing center.

The Governor of Illinois and the University of Illinois have contributed to the building of this center some $21.5 million over the five years of this program. In addition, Cray Research is supplying some $8 million, and the National Science Foundation's share is $43.9 million, for a total of $73.4 million over five years. Thus, the State of Illinois and Cray Research are supporting 40% of the Center's program, with the NSF supporting the remaining 60%. This partnership between Federal, State and industrial sources together with the University is what is making the national centers possible.

However, to build the quality of center which we were selected for by the peer review process, the five-year commitment by the NSF must be met and sustained beyond the initial five years. These centers, if successful, should provide technological leadership well into the next century. Because of the concern with the Federal deficits, the first year and a half (through October of 1986) of our program, as well as that of the other centers, has had to be severely curtailed. We expect to make up for that in the later years of the proposal, but the Congress should be aware that a significant ramping-up of funds is still required for all the centers. Even if this happens, the Congress should be aware that the budgets which

the National Science Foundation has approved are bare-bones budgets. They will just barely enable us to set up a center and maintain it at the state-of-the-art, but there is very little room for the kind of development of new software and hardware which we believe is crucial for the promise of these centers to be realized.

The National Science Foundation's Bardon-Curtis Report had originally envisioned a $100 million per year program as being necessary to meet the requirements of a national program. As a national center director, I certainly believe that is the correct order of magnitude, representing a factor of two greater expenditure than the current FY86 budget. It would be a tremendous waste of the funds already invested if the centers, once started, were starved for the necessary funds to carry out their missions.

I urge the Congress to consider carefully in its fiscal 1987 budget the out-year requirements of the national centers.

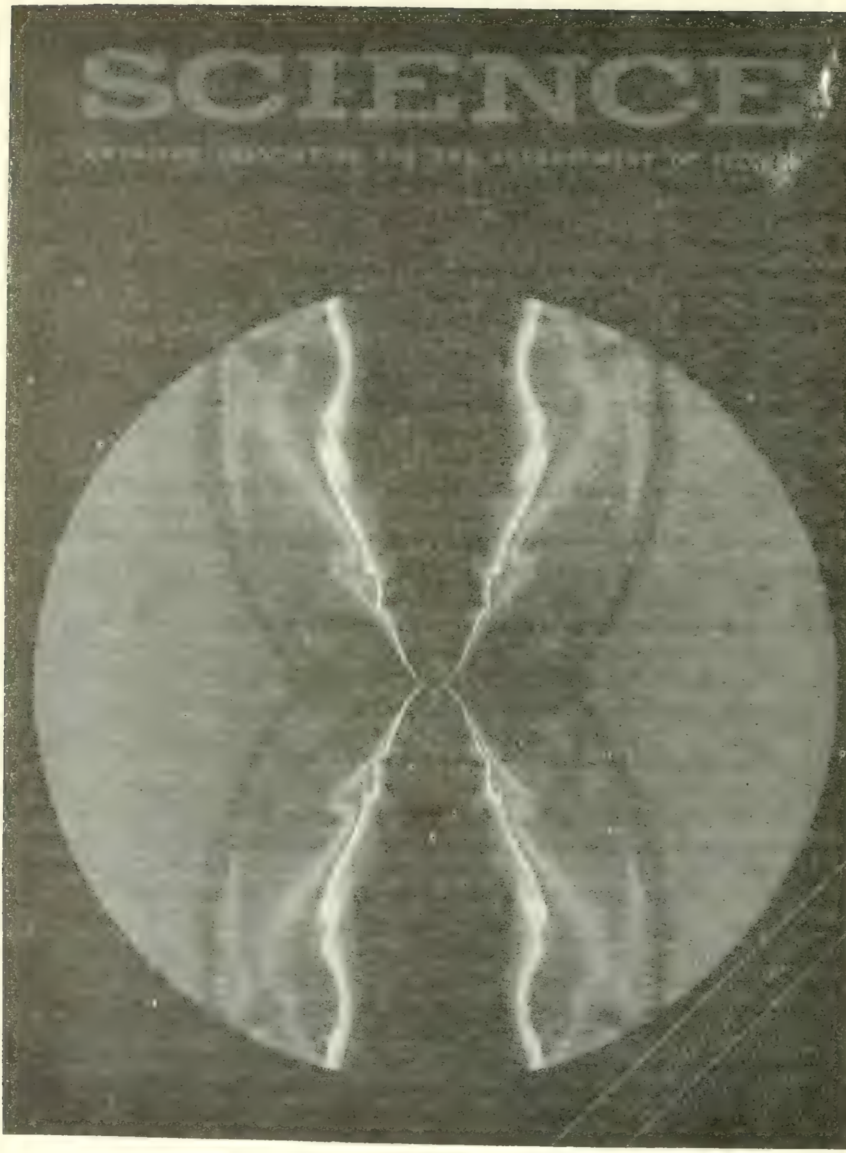
In addition to setting up the national centers, the national network must be built including workstation access and local networks within the univesities. Furthermore, as the enormous number of new users of supercomputers are generated by the national centers, the demand for supercomputer capacity is going to rapidly increase.

I think it is important to realize that there is an enormous economy-of-scale to be had by augmenting the computational resources at the established centers rather than by creating new centers.

To add a second supercomputer is a fairly marginal cost to an operational center; funds must be found for the second machine and the incremental staff needed to run it, but these costs are a small fraction of the cost of establishing a new center from the ground up. Therefore it is probably more sensible to follow a two-track goal to increase national supercomputer capacity. The first track would place more of the latest supercomputers in the existing centers as saturation occurs. The second track would place the new "mini-supercomputers", which cost under $1 million each, in remote university research groups which have demonstrated need for large amounts of production time.

In summary, Congress and the NSF have started a bold initiative which will have far reaching consequences for America's basic research effort. Furthermore, there is a promise for revitalizing American industry if corporations join with the partnership already formed between the Federal and State governments, universities, and the computer vendors. The key thing to remember is that this is a long term partnership which is going to require increasing investment for some years to come.

Assuming we can all keep the vision of coupling the computer and scientific revolutions, which brought us together, I am confident of undreamed of results from this program.


SCIENCE

26 April 1985, Volume 228, Number 4698

SCIENCE

An Approach to Complexity: Numerical Computations

Larry L. Smarr

Newton established modern mathematical physics in 1687 with the publication of his *Philosophiae Naturalis Principia Mathematica*, in which he showed how infinitesimal calculus could be used as the fundamental mathematical language of science. Since then, calculus has been instrumental in the discovery of the laws of electromagnetism, gas and fluid dynamics, statistical mechanics, and general relativity. These classical laws of nature have been described by partial differential equations (PDE's) for a continuum field.

Summary. The use of supercomputers and modern color-imaging techniques for numerical computation is beginning to fulfill von Neumann's vision that digital computers would become the most appropriate tool for solving nonlinear partial differential equations. An example of this approach, a model for the gas flow in the vicinity of a black hole, is described. From such calculations comes a realization that the multidimensional, dynamic solutions of nonlinear partial differential equations can exhibit complex behavior compared to what one normally encounters in analytic solutions. This complexity includes small-scale chaotic structure and large-scale persistently ordered structure. Computational methodology and the aesthetics that derive from it are discussed.

The tools of calculus have also proved powerful for discovering exact analytic solutions of these equations. Indeed, for linear PDE's that are separable, and thus reduced to ordinary differential equations, such techniques as Fourier analysis can give all solutions of the equations. Much of theoretical physics is based on the results of such linear analysis. However, progress in solving the nonlinear PDE's that govern a great portion of the phenomena of nature has been less rapid. Here, exact analytic methods and linear perturbation methods have provided only a tiny subset of the set of all possible solutions (the solution space) of the equations. There are probably vast regions of the solution space in which the character of the solutions is qualitatively different from the character of the known analytic solutions.

The author is director of the National Center for Supercomputing Applications and associate professor in the Departments of Physics and of Astronomy at the University of Illinois, Urbana 61801

An alternative approach is to find approximate but general solutions for the nonlinear PDE's by the use of finite differences instead of infinitesimal differentials. In this approach, the space-time continuum is replaced by a discrete space-time lattice of events, and the PDE's are converted into a large set of coupled algebraic equations. The unknowns in the algebraic equations represent the field's values at each point of the lattice. With enough time, a computer can solve the algebraic system for the discrete solution. In principle, no symmetry or time independence need be imposed. As the spacing of the lattice is made smaller, the discrete solution should approach the one for the continuum.

Modern supercomputers—the computers with the largest memories and the fastest processors—are making this alternative approach quite practical. The result is a revolution in our understanding of the complexity and variety inherent in the laws of nature. Not surprisingly, these more realistic solutions are allowing a more constructive interplay between theory and experiment or observation than has heretofore been possible.

All this was foreseen by John von Neumann, who I believe occupies a position similar to that of Newton. The mathematician Garrett Birkoff makes this point strongly in a recent review article about numerical fluid dynamics. He paraphrases von Neumann's vision as follows (*1*):

> It seems clear . . . that von Neumann was envisioning fluid dynamics as a *mathematical science* as had Euler, Lagrange, Stokes, Riemann, and Poincare before him. His main point was that mathematicians had nearly exhausted *analytical* methods, which apply mainly to *linear* differential equations and *special geometries*. . . . In short, von Neumann's proposal was that, with high speed digital computers, one could substitute *numerical* for analytical methods, tackling *nonlinear* problems in *general* geometries.

Birkhoff notes that since von Neumann made these remarks, computing machines have increased in speed by a factor of one billion and become cheaper per computation by a factor of ten million. Assessing where we are today, he concludes (*1*):

> . . . numerical fluid mechanics has not and will not replace either analytical or experimental fluid mechanics as a research tool, but . . . it complements and supplements them invaluably.

Extending this view from a comment on fluid mechanics to a general conclusion, the mathematician James Glimm recently wrote (*2*):

> Computers will affect science and technology at least as profoundly as did the invention of calculus. The reasons are the same. As with calculus, computers have increased and will increase enormously the range of solvable problems. The full development of these events will occupy decades and the rapid progress which we see currently is a strong sign that the impact of computing will be much greater in the future than it is today.

With my colleagues (*3*), I have been practicing the "von Neumann approach" for the past 10 years on a wide variety of physical problems. Out of this research has developed a well-defined methodology for attacking a broad range of problems that occur in physics. Our approach to complexity complements other methods but has unique characteristics of its own. Many of the techniques and concepts that are described below were developed by M. L. Norman and K.-H. A. Winkler during our research on the modeling of supersonic gas jets (*4*). Since that research is well documented in the literature, I will illustrate the methodology with a more recent project I worked on with John Hawley (*5*).

This project explored what happens when gas falls toward a black hole. It is not my purpose to explain in detail the theory of black hole accretion; rather, I hope the description of this project will show the paradigmatic aspects of the methodology used in any computational approach to solving the PDE's that define the laws of physics. In addition, the black hole example exhibits those features of complexity that appear to be common to many solutions of these nonlinear dynamic PDE's, in particular, a large-scale persistently ordered structure, which imposes itself on the underlying gas flow. The coherent structure is spatially complicated and slowly changes with time, but the important point is that an approximate simplicity and certain morphological features appear at a higher level of complexity than might have been expected.

Numerical Modeling of Black Hole Accretion

Gas flowing toward black holes is believed to be the mechanism that drives the "central engine" in quasars and active galactic nuclei. Besides generating great luminosity in the vicinity of black holes, the gas flow may generate the bidirectional outflowing jets of gas that are often observed emerging from galactic centers (*6*). Because black holes are so small, direct observation of the gas dynamics around them is impossible. Therefore, our only hope for an understanding of this phenomenon is to solve the equations that govern it.

To this end, Hawley and I, in collaboration with James R. Wilson (*7*), developed a computer program that solves the general equations that describe relativistic gas dynamics in the fixed gravitational field of a black hole. The program requires the gas flow to be axisymmetric, but no other symmetry is imposed. The gas obeys the ideal gas law; however, shock discontinuities are allowed and are modeled by an artificial viscosity (*7*). No effects from real viscosity, heating or cooling, radiation, or magnetic fields are taken into account in the program. The nonlinear, coupled PDE's that describe the gas flow (*7*) closely resemble the standard Newtonian ones representing mass continuity, energy conservation, and the change in momentum caused by the effects of (relativistic) gravity. The solutions of these PDE's are the five functions needed to specify the density, the energy density (or, equivalently, the pressure), and the velocity fields on the static curved space-time continuum of the black hole.

Our goal was to discover what happens to a rotating gas flow as it falls toward a black hole. To this end, we performed a series of experiments numerically. We chose appropriate initial conditions and boundary values to represent such a gas flow, and then we used the finite-difference versions of the PDE's for relativistic gas dynamics to calculate the change in the gas flow with time. With modern color-imaging techniques, we could watch the flow develop as if we were watching an experiment in a laboratory.

We model space with a grid based on spherical coordinates. In the solutions exhibited here, there are 160 evenly spaced angular zones from the north pole to the south pole and 160 radial zones from the surface of the black hole to the outer spherical boundary (Fig. 1). The radial spacing is increased with distance from the hole to keep the grid zones approximately square. The PDE's are converted to finite-difference equations (FDE's) by replacing the differential operators in the PDE's with Eulerian finite-difference operators on the grid (*7*).

The resulting PDE's constitute a set of coupled, nonlinear algebraic equations. The unknowns are the $160 \times 160 =$ 25,600 values of the five physical variables at each time step. The computations are started at some instant of time by assigning all the unknowns initial values, and then the equations are solved for discrete steps in time. A typical experiment makes use of at least 10,000 time steps. Thus, the finite-difference solution is a set of five variables on a space-time lattice of 250 million points, that is, the solution is 1.25 billion numbers. One of the key issues I deal with in this article is how to translate this hopeless pile of numbers into recognizable science.

In the problem described here (*8*), we assumed that a new supply of gas had begun to fall toward a black hole that had previously consumed all the gas in its vicinity. To specify the rotation law for the gas, we assumed that all the gas had the same value of specific angular momentum. The calculation was started at the time when the inner edge of the infalling gas had reached a boundary radius 50 times the black hole's diameter, and the gas was assigned a radial free-fall velocity appropriate to gas that has fallen from infinity. Thereafter gas continually poured in across the outer boundary. If the gas reached the surface of the hole, it was removed from the grid.

What was the final state of the gas flow? Before we made our calculations there had been little insight into the near-hole dynamics of nonspherical gas flow. However, a number of analytic studies (*7*) had given us clues to the character of any such flow. In the exactly soluble time-dependent, nonrotating, spherical case, the gas flow became supersonic before the gas fell into the hole. Therefore, as the flow became nonspherical in the general case, we expected that shock waves would become important.

In another exactly soluble case, for time-independent, nonspherical, rotating equilibrium, the natural state for a hot pressure-supported gas with constant specific angular momentum was an orbiting thick disk. Such a thick disk (Fig. 1) is shaped like a bagel with the black hole at the center of the bagel hole. The closed pressure-contour lines show how the pressure decreases with distance from a maximum near the surface of the black hole. Because constant specific angular momentum results in a vortex flow, the rotational velocity rises without limit as the axis is approached. Therefore, the centrifugal acceleration of the gas will always overcome the gravitational inward acceleration, resulting in an excluded funnel interior to this vortex flow. The static funnel wall threads through the opening in the bagel hole.

The first use of these analytic solutions was as calibrators for our program. We extensively tested (*7*) various differencing schemes in our program to determine which ones most accurately reproduced the analytic solutions. Second, the analytic solutions suggested features for us to look for in the general problem. For example, do shock fronts develop? However, analytic considerations could take us only so far. The detailed solution of the fully nonlinear, time-dependent, multidimensional, coupled PDE's was need-

ed to see whether the incoming, nearly spherical, supersonic flow could form a highly nonspherical, subsonic, orbiting thick disk.

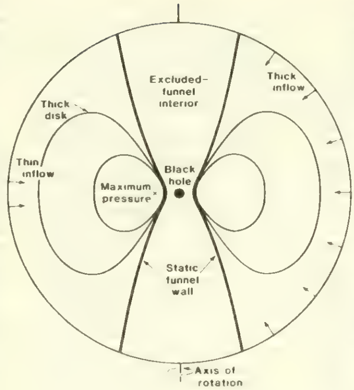


Fig. 1. A schematic diagram of the black hole accretion problem. The black hole in the center creates the gravitational field in which the gas flows. The gas is orbiting the hole with a constant value of specific angular momentum. The axis of rotation is vertical, and the equatorial plane is horizontal. The diagram is a cross section of the gas flow, and the contours of pressure for a thick disk at static equilibrium are shown. The pressure decreases outward from the pressure maximum near the hole. The static funnel wall is the closest to the axis the gas can come, given its angular momentum. Inside this wall the funnel is empty. There are two classes of boundary conditions: thin inflow, with gas only entering the grid near the equatorial plane, and thick inflow, with gas flowing in at all angles not excluded by the funnel interior.

Exploring the Solution Space

Having set up our experiment, we began to explore the properties of the solution space. The first step was to define the dimensionless parameters that span the space. For a given value of one parameter specifying the flow, the specific angular momentum, the flow will change with the variation of another parameter, the ratio of the solid angle of the incoming flow to the angle subtended by the funnel wall. In Fig. 1, the boundary conditions for the two extremes of thick and thin inflow are indicated.

Thus, with the initial conditions and these boundary values, we selected the solution space of the PDE's to be a two-parameter (thickness, angular momentum) family of gas flows in the fixed space-time of the black hole. Each computer run, which shows the development of the gas flow in time, is determined by one point in that two-parameter solution space. Our strategy was to spot-check the solution space by computing the results for both thick and thin inflows at a number of angular momenta.

The values of angular momentum were determined by the results of analytic theory. In Newtonian theory, any particle falling toward a $1/r$ gravitational potential finds a turning point at some radius. In general relativity, the gravitational pull increases faster than $1/r$; so for sufficiently small angular momenta, gravity overwhelms centrifugal acceleration, and the particle falls into the black hole, even though its Newtonian turning point would be outside the surface of the black hole (*7*). There is a critical value of specific angular momentum at which this effect first occurs. Therefore, for our computations we chose values of angular momenta that were more than, about, and less than this critical value, expecting qualitatively different behavior in the resulting solutions.

I will describe in detail only one computer run (*8*), which resulted from the choice of one point in the two-parameter solution space. My description will focus on the scientific results only insofar as they illustrate the scientific methodology used in this approach.

As gas pours onto the grid from the outer boundary, it begins to fall toward the hole (*9*). Figure 2 (*10*) represents a cross section of the density field at an early stage in the development of the flow, when the gas has just bounced off the centrifugal barrier near the hole. A single computer run would be represented by thousands of such images. The spectral order of colors creates a color scale (subdivided into 73 intervals) that we chose to be proportional to the logarithm of density, with blue representing the lowest value and red the highest value. In each image, each grid zone is assigned a color from this scale to represent the density of the gas in that zone at that moment. No color smoothing is done (the individual zones can be seen near the edge of the grid), so the color image accurately represents the results from the computation.

The central black hole and the evacuated funnel north and south of it are visible. From the outer boundary to smaller radii, the color gradually shifts from orange to redder colors, representing the adiabatic compression of the gas. There is a black, empty region surrounding the outside of the funnel wall bordered by thick red lines in an hourglass shape. The sudden jump from orange to red denotes a funnel-wall standoff shock front. This shock front is caused by the centrifugal deceleration of the incoming supersonic gas. It causes the gas to turn abruptly and slide down the inner edges of the standoff shock front.

When the two sliding gas flows (whose density has become so high that they are shown as dark red) meet at the equatorial plane, the gas is diverted inward and shoots toward the hole. As this rotating gas flow nears the hole, its centrifugal acceleration increases faster than the intense gravitational attraction of the black hole. At the last moment, the gas splashes backward off the centrifugal barrier. To avoid the continually incoming gas, the gas that splashes back must flow above and below the equatorial plane. During this process, some of the gas begins to build up a thick disk in precisely the region predicted by analytic theory. Remarkably, very little gas flows into the hole.

The color image on the cover shows the density field of the flow at a later time, after a quasi-steady state has been established; Fig. 3 shows the pressure field at that time. Because a strong shock front causes a much larger jump in pressure than in density, the color image of the pressure field is ideal for locating shock fronts (the color scale is proportional to the logarithm of the pressure). An abrupt jump from dark blue (lowest pressure) to bright blue or green across the funnel-wall standoff shock front is shown. Figures 2 and 3 also show that the shock front migrates to a larger radius with time. This is caused by the continual buildup of a high back pressure in the thick disk (the red heart-shaped region), which pushes the shock front outward.

In Fig. 3 the arrows represent the direction of the flow. Note that the shape of the thick disk is distorted from the equilibrium form shown in Fig. 1 by the ram pressure of the inflowing gas, just as might be expected intuitively. The in-

flowing gas is pushed away by the high pressure of the thick disk; however, it is trapped between the static funnel wall, where it is excluded by centrifugal acceleration, and the standoff wall shock. Its only means of escape is to flow out vertically as a hollow biconical jet. In doing so it adiabatically expands, and the pressure decreases, as shown by the colors changing from red to green to blue. The key features of the flow are the quasi-stationary patterns shown by the colors; the arrows show that gas continually flows through those patterns.

The question now becomes which of the features of the particular flow we have computed are generic. To answer this we have performed many computer runs, varying first one parameter and then the other (*8*). We have found that as long as the inflow is thick, the standoff shock fronts occur. As the angular momentum decreases, the general relativistic effect mentioned above opens a "spillway" from the inner edge of the thick disk into the hole [an effect predicted from analytic calculations by Paczynski (*11*)]. As more and more gas flows into the hole, the flow out of the jet is reduced. Finally, when the angular momentum is close to the critical value expected from general relativistic theory, all the gas inside the standoff shock fronts goes into the hole, and no thick disk or jet forms.

With thin inflow, a similar sequence of structures occurs in the thick-disk region. However, outside of the disk, no standoff shock fronts form and the narrow jet becomes a wide billowy wind. A black and white image of this configuration can be seen in a previous issue of *Science* (*12*). In the extreme case of thin inflow and very low angular momentum, the gas falls steadily into the hole.

In summary, we have found that the two-parameter solution space decomposes into regions within which the solutions share common morphological features. These features are not details of the flow but rather large-scale coherent patterns in it. For each distinct region of the solution space one can make a paradigmatic cartoon film of the solution. This procedure for characterizing a numerical function is not so different from what one does with analytic functions. Consider a sine function. One can draw a periodic, oscillatory, constant-amplitude graph to represent it without worrying about the particular values of the two parameters, period and amplitude. In both the analytic and numerical case, the important feature is the form of the function.

One of the key differences between numerical functions and simple analytic functions is that numerical functions have multidimensional spatial forms that are dynamic. That is, both Figs. 2 and 3 are frames from the same solution; the behavior in both figures must be included in the cartoon film representing this portion of solution space. The only way to understand the solution is to watch the color films that represent the solution in terms of different physical variables (*13*).

In summary, our approach is to compute discrete solutions to the finite-difference PDE's and then to convert these numbers into color images that change in time. In these images we can observe coherent large-scale structures in the flow. By performing additional computations, we can determine which of these structures are generic and over how large a region of solution space these structures are present. The boundaries between qualitatively different structures can be identified by this procedure. Finally, analytic methods and intuition are used to explain why the structures should be there. In some cases, this process reveals new phenomena for which complete analytic theories can be worked out.

I call this practical approach "exploring the phenomonology of solution space." This approach has a long history that was recently summarized by Zabusky (*14*). He terms the interplay between computing, analytic methods, and graphic visualization "computational synergetics." It is an important methodology of science and one that is becoming more widespread.

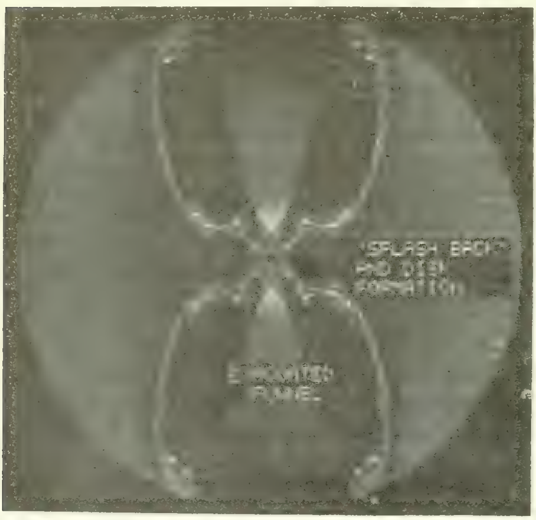


Fig. 2. The gas flow, in the thick-inflow case, at the moment the gas begins to splash back from the funnel wall near the hole. The quantity represented is density. The color scheme and features are explained in detail in the text.

The Ubiquity of Complexity

Just how prevalent is the phenomenon of complexity? Wolfram (*15*) gives an excellent overview of this question with particular emphasis on why the computer is so well matched to the study of complexity. It seems that most systems in nature can exhibit both simple and complex behavior. To date, theoretical physics has mostly concerned itself with simple behavior, since analytic tools were well matched to that study. However, as computational resources become more powerful and accessible, more studies are being performed on the complex behavior of simple systems.

A clear statement of the ubiquity of complexity, with examples from theoretical physics, can be found in "Prospectus for Computational Physics" (*16*), commonly referred to as the Press Report. This report shows that complexity can arise in a variety of ways: in moving from a few degrees of freedom to many degrees of freedom, from ordinary differential equations to PDE's, from simple or one-dimensional models of physical processes to multidimensional models, from low-order to high-order expansions, from scalar systems to vector or tensor systems, and from linear systems to nonlinear systems.

In summary, the Press Report concludes:

> One sees, then, that complexity arises not from 'bad taste in the choice of problems', but inevitably as theory advances . . . .

A similar statement could be made in every field of theoretical science. Ultimately, this is because nature is complex. Consider the dynamics and formation of galaxies, where a hundred billion stars interact gravitationally to produce the beautiful spiral arms familiar from astronomical photographs. At the opposite extreme of scale, the macromolecules that underlie life itself perform their biological functions largely because of the manner in which their thousands of component atoms arrange themselves in highly ordered large-scale structures. We are beginning to acquire the computational tools and scientific methodology that will allow us to attack these and other complexities head on.

Coexisting Aesthetics

Finally, let me turn to a disturbing feature of the revolution in computing techniques. Too often misunderstandings arise between scientists trained in classical analytic methods and those for whom numerical methods are the primary research tool. One often hears: "Numerical solutions are inelegant," or "Analytic solutions are simplistic." Such comments reveal a clash between two coexisting aesthetics derived from the nature of the computational tools that are used. Rather than define precisely the principles of both camps, I will give examples of their calculational goals.

Much scientific effort in the three centuries since Newton has been directed toward discovering the form of the basic laws of physics. Analytic methods have been precisely the right tool for that job. However, although analytic solutions have given us a skeletal view of the content of those laws, they have not revealed the true complexity of the solution space. Therefore, the search for form is shifting from the laws to the solutions of the equations describing those laws. Computational methods are the appropriate tools for this latter search.

Many of the classic analytic solutions are for fundamental static equilibria. What is becoming clear in many areas is that there is a new class of dynamic equilibria which are just as fundamental. These are large-scale coherent structures with long lifetimes compared to the underlying system's dynamical time scales. Although many are being studied observationally (for example, Jupiter's red spot), a growing number are being discovered numerically. As is the case for the soliton (*14*), the prototype of dynamic equilibria, for most of these structures some underlying mathematical principle is at work. It seems to be a general property of nonlinear systems that they "lock on" to coherent structures that are far from the linear regime.

Many analytic solutions exhibit high degrees of spatial or internal symmetry. Indeed the power of group theory in science attests to symmetry being a fundamental property of nature. However, many of the phenomena of nature are inherently unsymmetrical and time-dependent. The beauty of the ever changing three-dimensional structure of clouds is surely as great as the beauty of a perfect crystal. To explore such phenomena as the clouds requires the ability, which numerical tools give us, to probe complexity.

Much of the beauty of analytic functions comes from their encoding what is visually beautiful. For example, the periodic and oscillatory nature of the sine function is better perceived by a graph of the function than by looking at its name. Just so, the eye can perceive fundamental properties of complex solutions by using color images or other computational devices in situations where closed analytic forms are impossible. Thus, the visual representation of mathematical functions may become the common bond between simple analytic functions and complex numerical ones.

As I have attempted to show in my example of a black hole, scientists need to have an intuition formed from both aesthetics. There is no inherent conflict

Fig. 3. The gas flow at a later time for the same conditions as in Fig. 2. The colors represent the pressure gradient. The arrows indicate the direction of the flow.

between these two views; both are useful for discovering parts of the whole. I hope that students today are being trained with equal emphasis on analytic and numerical methodologies.

Prospectus

In summary, the prodigious growth in computing power is ushering in new approaches to complexity in many areas of science. Although the shift of methodology and aesthetics was foreseen by von Neumann over 30 years ago, the fulfillment of his vision is only beginning. For his vision to be realized, there are two major requirements. First, computers must continue their rapid rate of increase in speed so that more and more complex problems can be attacked on human time scales. Second, there must be much greater accessibility to the full range of computational tools that are needed so that a "critical mass" of scientists can work in each field of interest. Both of these requirements are likely to be met.

References and Notes

1. G. Birkoff, *SIAM Rev.* **25**, 1 (1983).
2. J. Glimm, in *Frontiers of Supercomputing*, N. Metropolis *et al.*, Eds. (Univ. of California Press, Berkeley, in press).
3. These colleagues include J. Centrella, B. DeWitt, P. Dykema, K. Eppley, C. Evans, J. Hawley, J. LeBlanc, M. Norman, J. Wilson, and K.-H. A. Winkler.
4. M. L. Norman, L. Smarr, K.-H. A. Winkler, M. D. Smith, *Astron Astrophys.* **113**, 285 (1982); M. L. Norman, K.-H. A. Winkler, L. Smarr, in *Astrophysical Jets*, A. Ferrari and A. G. Pacholczyk, Eds. (Reidel, Dordrecht, 1982), p. 227; M. L. Norman, L. L. Smarr, K.-H. A. Winkler, in *Numerical Astrophysics: A Festschrift in Honor of James R. Wilson*, J. Centrella, J. LeBlanc, R. Bowers, Eds. (Jones and Bartlet, Portota Valley, Calif., 1985), p. 88; L. L. Smarr, M. L. Norman, K.-H. A. Winkler, *Physica D* **12**, 83 (1984); M. L. Norman, K.-H. A. Winkler, L. L. Smarr, in *Physics of Energy Transport in Extragalactic Radio Sources*, A. Bridle and J. Eilek, Eds., in preparation.
5. This work was performed while J. Hawley was a graduate student in the Astronomy Department at the University of Illinois. He is currently the Bantrell Fellow at the California Institute of Technology.
6. M. C. Begelman, R. D. Blandford, M. J. Rees, *Rev. Mod Phys* **56**, 255 (1984).
7. J. R. Wilson at the Lawrence Livermore National Laboratory was the first person to write a computer program that solved the general relativistic hydrodynamics equations for this problem, see J. R Wilson, *Astrophys. J.* **173**, 431 (1972). Hawley, Wilson, and I wrote an updated version of the program and carefully calibrated it; see J. Hawley, L. L. Smarr, J. R. Wilson, *Astrophys. J.* **277**, 296 (1984); *Astrophys. J. Suppl. Ser.* **55**, 211 (1984)
8. J. Hawley, "A numerical study of nonspherical black hole accretion" [thesis, University of Illinois (1984)]; J. F. Hawley and L. L. Smarr, *AIP Conf. Proc.*, in press. Only black and white images are in these two sources.
9. In this discussion we assume that the black hole is nonrotating. Our program also handles the case of a rotating black hole. See (*8*) for details.
10. K.-H. A. Winkler developed the color-image algorithm for presenting the supercomputer solutions. J. Hawley adapted the approach and rewrote the algorithm for our project. The calculations were made on the Cray-1 supercomputer at the Max-Planck Institute for Physics and Astrophysics, Institute for Astrophysics.
11. B. Paczynski, *Astron. Astrophys.* **34**, 161 (1974).
12. M. M. Waldrop, *Science* **223**, 575 (1984).
13. The movie was made at the Max-Planck Institute for Physics and Astrophysics, Institute for Astrophysics.
14. N. J. Zabusky, *Phys. Today* **37**, 36 (July 1984); *J. Comp. Phys.* **43**, 195 (1981).
15. S. Wolfram, *Sci. Am.* **251**, 188 (September 1984).
16. W. H. Press, "Report by the Subcommittee on Computational Facilities for Theoretical Research to the Advisory Committee for Physics, Division of Physics" (National Science Foundation, Washington, D.C., 1981).

17. I thank J. Hawley for permission to use the unpublished color images in this article. My colleagues have been instrumental in the development of the ideas expressed in this article. In particular, I thank J. Centrella, C. Evans, J. Hawley, M. Norman, W. Press, R. Wilhelmson, K. Wilson, J. Wilson, K.-H. A. Winkler, and N. Zabusky. The hospitality and support of the Max-Planck Institute for Physics and Astrophysics, Institute for Astrophysics, and its director, R. Kippenhahn, are greatly appreciated. The images were made both at the Institute for Astrophysics and at the University of Illinois VAX and Image Processing Center. This work was partially supported by the National Science Foundation and the Alfred P. Sloan Foundation.

COVER

The color image represents a cross section of a spiraling gas flow near a black hole. The colors represent the gas density in every zone on the computational grid; density decreases in spectral order from red to blue. The underlying grid can be seen as the small squares at the outer edges of the image. The red stripes that "pinch in" at the equator (horizontal plane) are shock waves. The heart-shaped region is a fat disk orbiting the black hole. See page 403. [Computation by John Hawley and Larry L. Smarr, Department of Astronomy, University of Illinois. Calculations were performed on the Cray-1 supercomputer, Max-Planck Institut für Physik und Astrophysik, Garching, Federal Republic of Germany]

Mr. FUQUA. Thank you very much.
Dr. Wilson.

STATEMENT OF DR. KENNETH G. WILSON, JAMES A. WEEKS PROFESSOR OF PHYSICAL SCIENCE, NEWMAN LABORATORY; AND DIRECTOR, CENTER FOR THEORY AND SIMULATION, CORNELL UNIVERSITY, ITHACA, NY

Mr. WILSON. It is my great pleasure to be here to be welcomed to the State of Florida by the chairman of our Committee, Don Fuqua.

I hope at some future date Representative Sherwood Boehlert and I will be able to welcome all of you to upstate New York, and I can assure you that if you come, you will feel that New York State has abolished air transportation. [Laugther.]

I would like very much to thank the Congress and especially this committee for their work and support over many years on the supercomputer problem, and in particular, the funding levels for 1985 of $40 million, largely through the efforts of this committee made the award of a supercomputer center at Cornell University possible.

Without that award, you would not see me here. And also the diversity which the Cornell award has provided to the whole NSF Program. And so I am very grateful to you on my personal behalf, on behalf of Cornell and on behalf of the NSF Program, to you, Don Fuqua, and to you, Sherwood Boehlert, and other members of the committee.

I would also like to thank the National Science Foundation, especially, for the very careful review process which led to the awards which meant as a byproduct we got a lot of expert advice on how to finalize the details of the plans for our center, and I can assure you from my personal experience that embarking on one of these supercomputer center projects was enormously complex, difficult and having expert advice from the review process was just invaluable.

As Larry Smarr mentioned, all the supercomputer centers have substantial support from outside of the Federal Government and I would like to take this opportunity to thank our other principal partners in the Cornell venture; that is, IBM, which is contributing over $30 million in terms of goods and services to our program; Floating Point Systems of Portland, Oregon, which is contributing discounts on computers and a level of technology which we could not obtain from any other source; and the State of New York, which is a very enthusiastic partner in the Cornell effort.

And I should mention that the State of New York has promised to embark on a major networking effort within the State that I hope will be part of the general Federal, national networking effort.

Now we have been discussing some of today's problems that require supercomputing. You had an especially eloquent talk from the representative of the Ford Motor Company earlier. But there are also problems for tomorrow, problems which really cannot be solved with today's supercomputers.

For example, there is the whole subject of—for example, electrons which move in very irregular paths inside atoms and mole-

cules, and without detailed understanding of those electrons, you cannot take full advantage of our basic understanding of the molecular structure of an atom.

One area which is of utmost importance to the future economy of this country is the question of new materials; new materials for manufacturing, new drugs for medicine.

It is hard for me to describe the diversity of materials that could appear, except to remind you that whole industries today are built on particularly simple material; aluminum, steel, oil, glass.

And yet we know from materials like carbon fibers, superconductors and so forth, that the discoveries of the future will dwarf anything that we know today in terms of the diversity of material on which our industry is based.

And yet to reach that diversity that we should be able to achieve in the future, we must have better understanding at a very fundamental level of how atoms build into molecules and how molecules build into the materials that we will achieve in the future.

And yet the understanding of that problem is a problem that we cannot really achieve with today's computers, it is just so much more difficult than the problems that are being done with today's computers, we must have tomorrow's computers to get into that area in a very substantial way.

Now as part of tomorrow's technology, one thing I would like to emphasize is that supercomputers cannot be restricted by their high cost to just a few national laboratories, to a few university centers, to a few top Fortune 500 companies.

Supercomputers of tomorrow will have to be mass produced at a reasonable cost. The entry level has to be below $100,000 so that literally thousands of corporations, tens of thousands of engineering and scientific teams can have their own supercomputer, fully dedicated to their own needs.

And I believe that the computing technology that is available and is becoming available makes that possible and an extraordinary transformation in the whole nature of supercomputers so that instead of being restricted to a few favored people such as might appear at such a meeting here 2 years ago, that a meeting 5 years from now addressing supercomputers would have to be held in the Houston Astrodome.

Mr. Gore. Or the Florida State Stadium. [Laughter.]

Dr. Wilson. I hope they made it big enough.

The 1990 supercomputers will serve to provide a comprehensive productivity advance throughout industry, throughout universities, throughout the Government laboratories.

Cornell's program is uniquely dedicated to moving toward that future as soon as possible. We are already up and running in terms of an environment which is today's supercomputing environment rather than tomorrow's, but we have been working with industry on a confidential basis for some time to try to make sure that tomorrow's technology will be here on Cornell's campus in the very near future.

We are working not only with respect to the supercomputers themselves, but with the entire supercomputing environment, what we call 1990 computing environment. And we have done this in

part because Cornell has an extraordinary strength in terms of its people and past experience to try and move towards the future.

I would mention most particularly Don Greenberg, who is a world leader in computer graphics, which is absolutely critical to our efforts. My co-principal investigator is Ravi Sudan in plasma physics; Ken King, who is a director. And so Cornell is looking to the future, and we are doing it with the very strong support of our partners, Floating Point Systems and the IBM.

Our major gap in the program at Cornell at the moment is that we have not fully signed up the number of major industrial users of computers who we believe have to play a fully equal role in the computing industry with universities in determining the future technology and making sure that it will meet their needs, which will be vastly larger than the needs that were decribed earlier today by Ford Motor Company.

I would like to say we would welcome any further discussion from industries represented here today on the nature of our program and how they might participate.

Now I would like to comment on networking. I think it is the most important issue for us to be addressing at the present time.

First of all, networks, once in place, are permanent. The computers that exist on the networks get replaced every few years as better computers come along. It is like replacing the bridges on the highways and the highway systems themselves are permanent—the networks are also permanent.

For the purposes of the supercomputing program, those networks do not yet exist and any help the Congress can give in making sure they do come into existence in the near future would be very welcome.

There is not nearly enough on the NSF programs to build a proper national network, but there is also not a critical mass of people inside NSF to make that network come into existence.

There must be cooperation with the Department of Defense, with NASA and the Department of Energy. And this committee, I hope, will take a personal responsibility among all its members to keep the pressure on all those agencies to put enough people together to have the critical mass to make that network come into existence and not be subject to the usual interagency confusion and lack of cooperation.

Of course, building a network will require the participation of states and the private sector. It is really a major task, as Larry Smarr has said, but there is probably no more important task that this committee can be thinking of.

[The prepared statement of Mr. Wilson follows:]

Statement

by

Kenneth G. Wilson
Newman Laboratory, Cornell University
Ithaca, N.Y. 14853

Delivered to the House Committee on Science and Technology,
in session in Tallahassee, Florida

I am very happy to be asked to report to you on developments in supercomputing and the Advanced Scientific Computing Initiative of the National Science Foundation.

First, I would like to thank the Congress, and especially the Members of the House Committee on Science and Technology, for the strong support given to the Advanced Scientific Computing Initiative. Without your support, the program of the Initiative would be far weaker and more narrowly focussed than it now is. More particularly, the award of a supercomputer center to Cornell could not have happened without your help.

I would also like to thank the National Science Foundation for the careful review process used to select the winners of their competition. The continued interaction with the National Science Foundation staff and its expert team of outside reviewers has been extremely valuable to us in preparing the details of our own necessarily complex supercomputing plan.

It is too early to see any major results from the Initiative. Cornell has been operating an interim supercomputing system since May 1, but our facilities are still primitive and error-prone; the full program at Cornell and at the other centers will not be underway until the fall or winter.

There are a number of general comments and concerns that I would like to mention.

High speed networking on a national scale remains critical to the success of the Initiative. The appointment of Dennis Jennings to head the networking effort at the National Science Foundation has laid the basis for progress in the networking area, but continued attention and funding will be needed for progress. High quality graphics displays are becoming essential tools for research using supercomputers; as such graphics displays become available to supercomputer users around the country, the pressure for high speed networking will grow very rapidly.

Supercomputers are unique instruments of science in their breadth - every scientific discipline, every engineering discipline, many areas of medicine, agriculture, social sciences, architecture, etc., require supercomputer access now or in the future. Funding for the Centers is very meager when this breadth is

considered. For example, Cornell University will receive roughly seven million dollars per year from the National Science Foundation to operate one of the four national centers. The total sponsored research budget for Cornell's graduate programs is over one hundred and seventy million dollars per year; seven million more dollars could easily be justified in order to meet Cornell's supercomputing needs alone. By comparison, a major telescope, an accelerator, or a magnet facility serves only one discipline, or at most, a very small fraction of the disciplines a supercomputer can serve.

Basic research is characterized by the search for far-reaching breakthroughs in understanding Nature which have many unexpected practical spin-offs and which also lay the basis for further research advances. This is especially true of supercomputing. Much of the research using supercomputers will be concerned with applying the basic Laws of Nature to real life problems. There are only a handful of these Laws, such as Newton's Laws of Mechanical Motion, Maxwell's Laws for electromagnetic phenomena and Schrödinger's equation for automatic and molecular phenomena. These Laws are very precise and as reliable as the rising and setting of the sun. They are applicable to an extraordinary variety of circumstances. For example, every metal, every biological molecule up to and including DNA, and every exotic material (such as carbon fibers), has properties completely determined by the Schrödinger equation. The problem facing researchers is how to unlock the secrets of these equations in countless practical situations. Supercomputers are a critically-needed tool for this research but breakthroughs in understanding how to use supercomputers will also be required. The most rewarding research breakthroughs will be those that lead to progress in many different application areas at once - from weather forecasting to drug design. Such breakthroughs will result from increased understanding of the Laws of Nature themselves and how to work with them effectively.

Researchers often pose sample problems to be solved, by themselves or by their graduate students, which illustrate basic features of the Laws of Nature but do not have much immediate practical value. It is the understanding which results from solving these problems that can have practical benefits. For example, when I was a graduate student I studied problems at the far-out fringes of Elementary Particle Physics; even my closest colleagues found it difficult to justify what I was doing. During this time, I applied for a Bell Laboratories Fellowship and stated on the application form that my research plans bore no relation whatsoever to the needs of the telephone company. I was lying. Although I had no way of knowing this at the time, ten years later my research had progressed in unexpected directions involving properties of matter which had become of very great interest to physicists at Bell Laboratories. Even today, almost thirty years and one Nobel Prize later, it is not possible to predict where the most important spin-offs of my research will occur: in oil exploration? in the design of new materials? or even in social or economic studies?

With this background, I want to emphasize the utmost importance at present of not targeting basic research to specific applications. We need basic research in all its generality leading to spin-offs throughout the civilian and military sector - to improve our economic competitiveness, to improve our health,

and to advance our defenses. More attention needs to be given to the spin-off process, but targeting basic research is not the answer. The Deputy Director of the Cornell Center, Ravi Sudan, is also Director of Cornell's Plasma Physics Center and he is now struggling with the cutbacks in the Plasma Fusion program; his research is basic and should be considered part of the U.S. basic research effort rather than as specifically tied to the fortunes of the thermonuclear program. Professor Sudan's background in Plasma Physics will be very helpful in dealing with many different kinds of fluid and gas flow problems that will be studied on Cornell's supercomputer. I think it is especially important that basic research not be tied politically to the success of specific development projects like the Strategic Defense Initiative. Instead, the U.S. should encourage basic research to pursue its traditional search for generality, and simultaneously improve the spin-off process so that important research breakthroughs are milked for all kinds of spin-offs rather than being locked into one development project.

One important trend that can improve the research spin-off process is the growth of the graduate engineering profession. Ph.D.-trained engineers are better prepared to help move university research results into practice than engineers with just an undergraduate degree. At Cornell, the engineering college has converted over the last twenty years from a predominantly undergraduate program to a balanced mix of undergraduate and graduate training.

It is important that U.S. civilian manufacturing industries be provided their fair share of access to Ph.D. talent. Success in international competition with its resulting effects on both the national debt and the trade deficit depends increasingly on using high technology, including supercomputers and university research results to design more competitive products for the world market. Ph.D. scientists and engineers are needed to help civilian industries take advantage of high technology opportunities to improve design and manufacturing. The 1990's could bring vastly more powerful supercomputers than any that are available today or are even talked about. The key to revolutionary advances is through the use of parallelism: computers which can execute thousands or even millions of operations simultaneously. Parallelism on the scale of thousands of operations has been amply demonstrated in an academic setting, notably at the University of Edinburgh in Scotland, which has two ICL Distributed Array Processors, each capable of executing four thousand operations simultaneously. The main need currently is to make available and win commercial success for highly parallel computing systems with computing power vastly beyond today's Cray, Cyber, and Japanese supercomputers. Cornell is pressing the U.S. computer industry to provide such a system to Cornell within the next year or two.

Each of the four supercomputer centers is likely to develop its own areas of special expertise. I cannot speak for the other centers, but I would like to comment on Cornell's specific role. Cornell's main focus is towards the future of supercomputing. It already has projects underway to address bottlenecks that could prevent such machines from being both academically and commercially successful. Cornell has projects in software productivity, very advanced graphics, parallel processing systems, and in key application areas for parallel processing.

The most important area of all for future very powerful systems is the area of atomic and molecular physics with applications to the properties and design of materials. All present supercomputers are hopelessly inadequate for solving such

problems; if one learns how to solve these problems on future machines, the payoff could be spectacular. Experimental physics, chemistry, and biology have not even scratched the surface of the totality of chemical and material substances that could be industrially important; breakthroughs in the understanding and design of materials could lead to whole new industrial groupings alongside the aluminum industry, silicon industry, glass industry, oil industry, etc., each of which has a single material as its base. Cornell is the principal center with management (namely me) with the training to conduct and assess research in this area and, in fact, I have a joint research effort with John Wilkins of Cornell trying to analyze and improve on some of the most promising methods for solving the Schrödinger equation for atoms and molecules.

At present the supercomputer market is small; the supercomputers are built largely by hand at enormous cost. Meanwhile, the need for supercomputing in private industry and universities is growing explosively. Unfortunately, the five million dollar entry-level price of today's supercomputers puts them out of reach of all but the largest corporations. Even in Fortune 500 companies the process of deciding to buy a multi-million dollar supercomputer requires three years of intense battle between groups of scientists and engineers who want it versus company accountants who demand, but do not get, a yearly payback schedule for the investment. Most industrial scientists and engineers prefer to avoid this battle and opt instead to buy much less powerful but much cheaper superminicomputers. The result of this has been twofold. Firstly, Digital Equipment Corporation has become the second largest computer company in the U.S. with several billion dollars a year in superminicomputer sales, while supercomputer sales are far lower, of order a couple hundred millions of dollars per year. The second consequence is that many U.S. industrial scientists and engineers are very poorly equipped to deal with today's intense international competition. Their computing jobs take days or weeks to run on their superminicomputers and they desperately need the faster turnaround that supercomputers could provide.

Through parallel computer design it is now feasible to produce new supercomputer product lines that would have an entry-level price below one hundred thousand dollars and then would be upgradeable in hundred thousand dollar increments all the way to hundred million dollar behemoths vastly suprassing today's supercomputers. Such supercomputers could be mass produced at low cost using today's advanced but mass produced VLSI silicon chips. Given proper support, they would enjoy a huge market from U.S. private industry which needs to migrate *en masse* from their present inadequate systems.

Unfortunately, the proviso "with proper support" is a key sticking point. Private industry has accumulated a massive amount of software targeted to its existing computers, none of which can easily be moved to the revolutionary new parallel systems. Such systems have to accumulate both systems support and new applications starting from scratch, with almost a twenty year lag between today's most mature computers. A massive effort will be required to overcome this twenty year lag. Cornell is trying to build a coalition of adequate size to deal with this problem, including many computing manufacturers, national laboratories, and about thirty leading corporate users of scientific computing. The key to our program is the coalition of industrial users who define the ultimate market for the new systems. It is taking time to build an Industrial Associates Program at Cornell, but we continue to work hard on it. Many universities have computer science projects related to parallel processing which will help our efforts; as far as we know, Cornell is the only university focussing on helping the commercialization of parallel processing so that its benefits will become very widely available to both universities and industry.

Mr. FUQUA. Thank you, Dr. Wilson.
Dr. Killeen.
[The biographical sketch of Dr. Killeen follows:]

Killeen, John: Director, National MFE Computer Center, Lawrence, Livermore National Laboratory University of California, Livermore, California.
Birth: July 28, 1925, in Guam.
Family: Wife: Marjorie W. (Lyman) Killeen; Children: Michael, Sean, Jack, Joan, Katherine, Ann.
Address: Residence: 1528 Campus Drive, Berkeley, California 94708. Business: Univ. of Calif. Lawrence Livermore National Laboratory, P.O. Box 5509, Livermore, CA 94550.
Education: AB, University of California, Berkeley, 1949, Physics MA Univeristy of California, Berkeley, 1951, Mathematics; PhD, University of California, Berkeley, 1955, Mathematics.
Positions: Present Position: Professor, Dept. of Applied Science, U.C. Davis and Director, National MFE Computer Center, Lawrence Livermore National Laboratory, University of California. 1957–74, Theoretical Physics Group, Controlled Thermonuclear Research Division, Lawrence Livermore Laboratory. 1956, Research Associate, Courant Institute of Mathematical Sciences, New York University. 1950–1955, Mathematician, Lawrence Berkeley Laboratory, University of California.
Military: U.S. Navy 1943–46.
Publications: 75 articles.
Research: Computer Applications in Controlled Fusion Research; Computational Physics; Equilibria, Stability, and Transport of Plasmas confined by Magnetic Fields.
Memberships: Sigma Chi Iota; Fellow, American Physical Society; Member, American Mathematical Society; Member, American Assoc. for the Advancement of Science.
Special positions: Editor, *Journal of Computational* Physics; Assoc. Editor, *Computer Physics Communications*; Editor, *Methods in Computational Physics,* vol. 16; Editor, Springer Series in Computational Physics.
Awards: USDOE, Distinguished Associate Award, 1980.

STATEMENT OF DR. JOHN KILLEEN, DIRECTOR, NATIONAL MAGNETIC FUSION ENERGY [MFE] COMPUTER CENTER, LAWRENCE LIVERMORE NATIONAL LABORATORY, LIVERMORE, CA

Dr. KILLEEN. I would like to thank the committee for inviting me to tell you about the National Magnetic Fusion Energy Computer Center and our needs and plans.

This center is operated for the Department of Energy at the Lawrence Livermore National Laboratory. And, as Al Trivelpiece described earlier, it was conceived in 1973 in response to the recognition that the fusion program—magnetic fusion program—would require a considerable build-up on computational capability in order to supplement the experimental programs.

The fusion program is characterized by rather large experiments nowadays, large Tokamaks, and tandem mirrors and so on. And there are certain physics issues which need to be addressed in these experiments, but also an extensive program of computer modeling. Some of these are equilibrium and stability and perpendicular confinement, heating of the plasma to thermonuclear temperatures, fueling and so on. And in order to address these issues on a computational basis, a wide variety of models are needed.

The reason for this is the time scale of the plasma phenomena go over orders of magnitude from milliseconds all the way up to seconds, and the spatial scales also go over orders of magnitude from the size of the device down to things that are occurring on a microscopic basis. So, various three-dimensional fluid codes and trans-

port codes, equilibria codes, are needed to supplement the design of experiments as well as kinetic theory codes.

In my written testimony, which I would like to have made part of the record, I describe in some detail three of these models—time-dependent magnetohydrodynamics and particle models and Fokker-Planck models and their need for the most advanced supercomputers available.

The center, our center, was established in 1974. We started operation in 1975 with a 7600. We also, at that time, formed a network which connected the major sites. There is a map in my report. The major sites were the Princeton Plasma Physics Laboratory, Oak Ridge, Los Alamos, and GA at San Diego.

At that time, we leased 50 kilobit per second lines from the telephone company and connected those sites to the computing center at Livermore. By 1982, all of those links had been replaced by satellite, and we now have a dual satellite. These are—when I say dual satellite, these are each 56 kilobits per second, and we are using two satellites. This was originally put in for reliability, but we actually use both data channels. So we have a total of 112 kilobits per second going to the sites, and the Princeton link is being upgraded. Each one of them is being upgraded.

Then this network has evolved over the years, and connected to each of those nodes there are a great many other smaller places, some of them not quite so small; like MIT is connected to Princeton with a 50 kilobit land line. And we also have 9,600 baud lines coming in from nearby universities. Coming into Princeton, for example, are NYU and Cornell and Columbia and Maryland and coming into Oak Ridge are Wisconsin and Illinois. And these are—there are a great many universities—in connection with an earlier question—that are using our center. But they are people that are, in fact, funded by the fusion program or other energy research programs.

There are other ways of getting into the center besides our own network in that we have gateways to the ARPANET, which you heard about earlier, and to TYMNET, which is a commercial network. Those are both dial-up connections. So, our network is in existence. It has been in existence now since 1976, and it continues to grow.

The computers at the center have also continued to grow, and in 1978—well, long before that—our 7600 was saturated. And, by that time, we also were aware of vector computers being available. So we obtained one of the earlier Cray 1's in 1978. This had a dramatic impact on the fusion program.

Some of these models I described in my written testimony were really not possible on the 7600. Three-dimensional resistive magnetohydrodynamics, for example—which explains major disruptions in tokamaks—was not really explained until we got the first Cray 1. And other—the neutral beam heating of tokamaks, such as done in the TFTR, for example, require rather large modeling efforts.

In 1981, we added a second Cray 1. This was mainly because of the needed capacity requirements. And in 1982, we began to study what our requirements were for the future, and we put together a report which came out as a DOE report at that time. We were working with the "Blue Book." And it went through many of these

models I discussed and pointed out that we really need machines more capable than the Cray 1.

So we put together a program to acquire the next generation of computer as soon as it was available. And two weeks ago, the first Cray 2 computer was delivered to our center. This is a four-processor computer with 64 million words of memory, as compared with the Cray 1 which has, say, 2 million words of memory. So we have the first Cray 2, and this machine was delivered on the 28th of May. And on the 5th of June, it was turned over to us by the Cray engineers. And by Friday of last week, the 7th, we were already running programs on it.

Now this, of course, sounds a little bit far-fetched, but we had a single processor version of the Cray 2 since before Christmas, which had 16 million words of memory. So we got our operating system—we have the same operating system on all our computers. It is called CTSS, and it is a time-sharing system, and this is the preferred way of computing in the fusion and energy research community.

In order to bring this up, we had to develop our own compiler for the Cray 2, and we did this. And also, various application codes were, in fact, running on this single processor version of the Cray 2 by the time we had the actual Cray 2 installed. And we feel this is going to have a dramatic impact, just as the first Cray 1 in 1978 did, on our ability to do all these dimensional problems.

As Al Trivelpiece pointed out earlier, 2 years ago we began to allocate 5 percent of the two Cray's that we had at that time to other energy research programs; in particular, high energy physics, nuclear physics, material sciences, chemical sciences, and so on. And, of course, that really just whetted the appetite of those groups. But that 5 percent enabled a number of groups to get codes running on the Cray 1.

In November of 1984, the beginning of this fiscal year, we installed the Cray XMP-22 for the energy research programs, and that XMP-22 is like two Cray 1's. It is two processors, sharing a 2-million word memory, and it is fully utilized now. And, as Al pointed out earlier, there are rather large allocations in high energy physics, materials science and some other things that Dr. Wilson just talked about. People are doing large-scale theory calculations, they are using the XMP in studies of the design of the Superconductor Collider, for example, and they are doing materials studies, Monte Carlo and materials.

And our plan is to go to the—well, to request money for fiscal year 1987 to replace that machine. That is a leased machine, an interim machine, and we would replace that machine with a class VII computer.

In the report, I have given a table of various computers that are presently available, and some that we think will be available in this timeframe.

One of the earlier questions had to do with Japanese computers, and there are three Japanese computers in the table: one from Fujitsu, Hitachi, and NEC. But more importantly, of course, the Cyber-205 and the Cray XMP that are currently being sold in rather large numbers in this country, of course; and then there is the Cray 2, which we have just taken delivery of. And there is a

Cray 3, a 16-processor, and an ETA GF10 computer being developed. There are also successors to the XMP. So we feel that in the 1987 timeframe, we will have a fairly good choice of an advanced computer to meet the needs of these programs.

This also involves expanding the network, which we in effect have already started doing—and Al Trivelpiece mentioned that earlier. Places like SLAC and Fermilab and Brookhaven have been added to the network, as has Florida State University.

The number of users at some of our sites, of course, has gone up considerably. LBL has a lot of people in high energy and nuclear physics, of course, and basic energy sciences, and so does Argonne. We have just put in a leased—well, it is leased, of course—we put in a land line to Argonne, 56 kilobit land line. To get another satellite link would take quite a long time; whereas, we could get what is called a BBS line to Argonne, and that is connected in turn to Fermilab with another 56 kilobit line. So we are expanding the network that exists already to take care of the energy research program, and we are putting in remote user service stations at a number of new energy research sites.

Our needs for the future basically are, as I indicated in the written testimony, to really have the availability of the most advanced supercomputers available. And we know how to use them, we have the models, we know how the codes that can be developed to use these—we have developed multitasking, which is a step that one needs to have in the program in order to use a multiprocessor. We have multitasking as part of our operating system.

As I mentioned earlier, the next generation of computers will have 8 vector processors and 16, sharing very large memories. I think we are talking now up to 256 million words of memory. And we are quite excited about tackling some of these fusion problems that we know that we can do with such a computer.

Another area that we are crucially concerned with in addition to the network, which I think we have heard quite a bit about, is adequate file storage. When you run it through a computer center of this kind and you have all of your users—and we have over 3,000 users—people do not send all their files back and forth over the network all the time. That is why we are able to get along with 56 kilobits per second, rather than megabits. But they have to keep their data and their codes at the center in a file storage media.

This is getting to be one of our major bottlenecks, just storing all the information. Codes typically do not run through to completion. Many times the problems will run and then they will take the problem off and then restart it at some later time, so you have a large amount of data that has to be kept, say, from one day to the next. And this is one of our major problems.

So I appeal to the centers in the audience to start thinking about file storage problems as well.

Thank you.

[The prepared statement of Dr. Killeen follows:]

THE ROLE OF SUPERCOMPUTERS IN MAGNETIC FUSION AND ENERGY RESEARCH PROGRAMS

John Killeen
National Magnetic Fusion Energy Computer Center
Lawrence Livermore National Laboratory
Livermore, CA 94550

Abstract

The importance of computer modeling in magnetic fusion (MFE) and energy research (ER) programs is discussed. The need for the most advanced supercomputers is described, and the role of the National Magnetic Fusion Energy Computer Center in meeting these needs is explained.

MAGNETIC FUSION

INTRODUCTION

During the early 1970's the U.S. magnetic fusion program supported at least fifteen varieties of experimental concepts. These were rather small experiments as compared to today's large facilities. During the years 1974 to 1980, the program went through a period of dramatic growth, but at the same time evaluations and reviews reduced the number of experimental concepts supported to the following six:

- Tokamak
- Tandem Mirror
- Reverse Field Pinch
- Stellarator
- Compact Toroids
- Elmo Bumpy Torus

The most advanced of the above concepts is the Tokamak, and all four of the major international groups have commissioned large facilities to establish the scientific feasibility of fusion.

All of the international groups are designing forms of "The Next Step," which is an ignition Tokamak. A common feature of these designs is their large projected cost, so at the present time none have been authorized.

A large tandem mirror experiment (MFTF-B) is being built in the U.S. Advanced stellarators are planned for Kyoto, Japan and Garching, FRG, and a large RFP is planned for Padua, Italy. New stellarator, RFP, and compact toroid facilities are being proposed in the U.S.

In all of these concepts, there are eight fusion physics issues which must be addressed as a complete plasma system, i.e., they are interdependent. They are:

MHD Equilibrium and Stability
Perpendicular Ion and Electron Confinement
Parallel Confinement
Electric Potential
Heating
Fueling
Impurity Influx
Alpha Particle Heating

In order to resolve these issues, i.e. to reach a state where a fusion reactor is feasible, the experimental programs must be augmented by a program of computer simulation to aid in the design and interpretation of the experiments and implementation of theory. The following plasma physics models are of importance to the fusion program.

Time-dependent magnetohydrodynamics
Plasma transport in a magnetic field
MHD and guiding-center equilibria
MHD stability of confinement systems
Vlasov and particle models
Multi-species Fokker-Planck codes
Hybrid codes

The need for such a variety of models is caused by the great variation in time and space scales[1] present in the plasma phenomena relevant to the eight fusion physics issues. The implementation of these models requires the most advanced

supercomputers available. The impact of new supercomputers on some of the types of models will be discussed later in this paper.

In addition to plasma physics models, advanced engineering computations must be made. Engineering models needed in fusion reactor design studies include:

Plasma engineering-burning plasma dynamics
Nucleonics
Mechanical design
Magnetic field analysis
Systems studies
Thermal hydraulics
Tritium handling
Safety and environmental studies

NEW SUPERCOMPUTERS

Supercomputers are the most powerful general-purpose computers available for information processing. Currently, supercomputers have the capability of performing hundreds of millions of arithmetic or floating point operations per second (MFLOPS) and are used in two general areas: real-time applications such as signal processing and scientific computing. In the race to build the next generation of supercomputers, scientists are experimenting with a variety of architectural designs. The new architectures will have as few as two processors with shared memories to extensive parallel architectures with hundreds of local memories and processors, all executing instructions simultaneously.

There are three types of parallel architecture capable of increasing performance by a hundredfold over today's state-of-the-art supercomputers. They are:

Lockstep vector processors,
Tightly coupled parallel processors,
Massively parallel devices.

When the execution unit operates simultaneously (in lockstep) on many data entities, the machine is said to have an array architecture. When the execution unit operates on sets of data, on an assembly line basis, the machine is termed a vector processor or pipeline processor. The CDC 205 and Cray 1 are examples of vector processors. The real beneficiaries of such vector processors have turned out to be multi-dimensional fluid codes, which are dominated by long vector loops.

A second architectural type capable of a hundredfold increase over state-of-the-art supercomputers is tightly coupled systems of a few high-performance processors. In principle, collaboration of these processors on a common task can produce the two orders of magnitude speedup that is needed.

The current trend in supercomputer architecture is toward tightly coupled systems with two to four vector processors typically sharing a large memory. Recent experiments suggest that these systems can be successfully used in parallel processing of scientific computations. The next logical step in this trend is toward systems with 8, 16, or more processors.

In the long term it is possible to build massively parallel systems, that is, systems with 1000 or more processors communicating with thousands of memories. In general, the scientist cannot manually find and manage parallelism for thousands of processors. Rather, the software must find it, map it onto the architecture, and manage it. Therein lies a formidable research issue for massively parallel computation.

The following two tables list (1) existing supercomputers, and (2) announced supercomputers. This tabulation employs only the few parameters usually contained in press-release-type information.

NATIONAL MFE COMPUTER CENTER

The MFE Computer Network (Figure 1) provides fusion researchers in the U.S. the full range of available computational power in the most efficient and cost effective manner. This is achieved by using a network of computers of

TABLE I Current Supercomputers

Organization	Fujitsu	Hitachi	CDC	CRAY	CRAY
Model	VP-200	S-810/20	205	X-MP/2	X-MP/4
announcement	Jul 1982	AUg 1982	Jun 1981	Aug 1982	Aug 1984
architecture (64 bit words)	vector (IBM compatible)	vector (IBM compatible)	vector	vector multi-processor 2 CPU	vector multi-processor 4 CPU
maximum performance (M FLOPS)	500	630	400	479	953
maximum main memory size (64 bit words)	32M MOS	32M MOS	16M MOS	4M Bipolar	8M Bipolar

TABLE II Supercomputers Now In Design

Organization	CRAY	CRAY	ETA	Denelcor	NEC
Model	2	3	GF10	HEP-2	SX-2
announcement (or project start)	1985	none officially	Sept 1983	May 1983	April 1983
availability	1985	1986	1986	1986	1985
architecture	vector multi-processor 4 CPU	vector multi-processor 16 CPU	vector multi-processor 8 CPU	scalar multi-processor 64 CPU	vector
maximum performance (M FLOPS)	1,000	10,000	10,000	4,000	1,300
maximum main memory size (64 bit words)	256M MOS	256M MOS	256M MOS	256M MOS	32M MOS

different capability tied together and to the users via dedicated data lines and dial up telephone lines. The concept of the NMFECC is that different levels of computer capability are provided at the various locations according to research priorities. At the national center (Figure 2), providing high level capability to the entire community, are two high-speed Cray 1 computers, and a Cray X-MP/2. Additional equipment at the national center includes processors and other ADP equipment for communications, file management, and data storage.

On May 28, 1985 the first Cray 2 computer system was delivered to the NMFECC. This computer has four vector processors and 64 million words of MOS memory. This system will give the fusion community the capability required for advanced plasma modeling as described in the next section.

At the next level of capability are User Service Centers (USC's): DEC-10 computer systems with direct high-speed access to the national center through PDP-11/40 remote communications control processors. There are now five operational USCs (Figure 1) in the field located at Princeton Plasma Physics Laboratory (PPPL), the Los Alamos National Laboratory (LANL), the Oak Ridge National Laboratory (ORNL), GA Technologies, Inc. (GA), and LLNL (for the mirror confinement program). A sixth USC, used in center operations, is located at the NMFECC itself.

A third level of capability is provided through the Network Access Port (NAP). MFECC designed the NAP to permit remote computers to be connected to the MFE network as remote hosts.

A fourth level of capability is provided by Remote User Service Stations (RUSS) at selected sites (Figure 1). RUSS stations provide users with the capability of printing output files locally on a 1000 line/minute printer and act as a terminal concentrator for up to 16 interactive terminal users. RUSS stations are connected to the nearest MFE-NETWORK communications processor over a 9600 baud dedicated line (Figure 1).

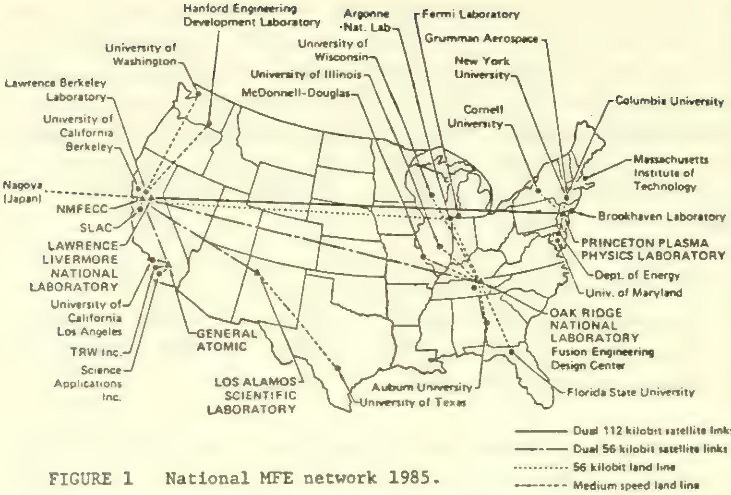


FIGURE 1 National MFE network 1985.

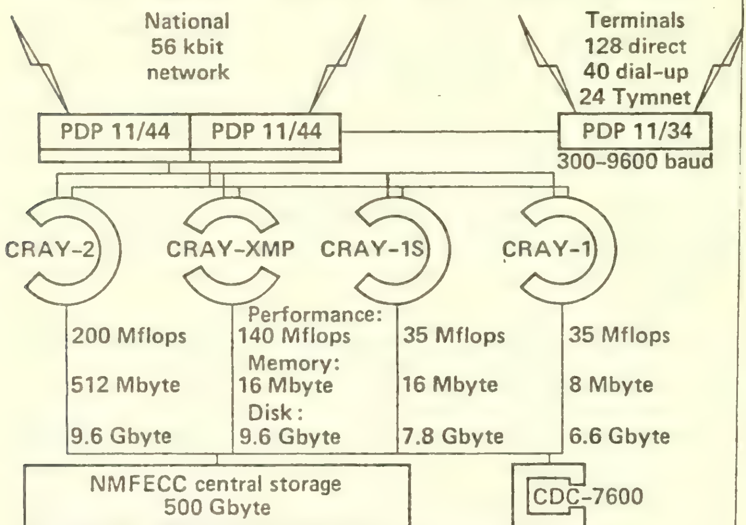


FIGURE 2 NMFECC hardware configuration.

Data-Communications Systems

Data Communications services to the National MFE Computer Center are provided on a 24 hours/7 day basis. Three types of service are provided to NMFECC users as outlined below:

1. Wide band Satellite Network Service. Users at Major USC's on the MFE net may log on to their local DEC-10 system and interact with the computing resources at the Central facility in Livermore.

2. Dedicated 9600 Baud Service. Remote User Service Stations on the MFE Net are served by dedicated leased 9600 baud lines which terminate either at the Center (LLNL) or at the nearest MFE Communications Control Processor (Figure 1).

3. Dial Up Service. Users not at major fusion laboratories may dial-up the Center using one of the following services: (a) TYMNET, (b) ARPANET and (c) DIRECT DIAL COMMERCIAL.

NMFECC Computing Environment

The NMFECC computing environment reflects the needs of computer users in the Magnetic Fusion Energy research community. Both interactive timesharing and batch processing are available. The fusion community has always found that interactive computing, even with the largest codes, is by far the most efficient use of physicists efforts. The 5% overhead in swapping codes in and out of the machines provides fast debugging, immediate turn around on key results, and the capability to interact with codes which need user control. The Livermore Time Sharing System (LTSS) was adapted by the NMFECC for the Cray 1 computer in about six months. CTSS is supported by libraries of FORTRAN callable subroutines which enable a user to issue almost every system call, giving access to every part of the hardware. A typical physics code can be run from a terminal, display graphics as it runs, be interrupted or interrogated at any time. The ability to start or stop a code at any point and inspect the results provides debugging at least 100 times faster than older methods. The CTSS operating system is also used on the Cray X-MP/2 and the Cray 2.

PLASMA MODELING IN MAGNETIC FUSION

It is within our grasp at present to model plasmas in full 3D with one or two orders of magnitude variation in space and time scales in each problem. Some of the recent success in the field are worth listing as they are the basis for further developments. Representative work in MHD, Kinetic models, and Fokker-Planck calculations are considered.

Time-dependent MHD Codes

A technique for determining MHD instabilities along with their growth rates is through the solution of the time dependent MHD equations of motion. The full set of MHD calculations comprise a coupled system of eight nonlinear partial differential equations, the solution of which is a formidable task on any computer system. In order to make these computations tractable, approximations have often been made, including reduction in dimensionality, linearization, restriction to a particular geometry, ordering, or regime, and the assumption of no transport or resisitivity.

The recent advances in three-dimensional resistive MHD calculations for tokamaks have depended crucially on obtaining a reduced set of MHD equations by expanding in the inverse aspect ratio[2]. This is possible because of the strong and almost uniform toroidal magnetic field in tokamaks. Additionally, the computational speed of the codes based on the tokamak reduced equations is greatly enhanced by the assumption of incompressibility, which eliminates the compressional Alfvén wave. Because of the strong field in a tokamak, the fastest remaining mode evolves on a time-scale on the order of the major circumference divided by the Alfvén velocity. This time scale may be more than an order of magnitude longer than that of the compressional Alfvén wave. Since the field components in the Reversed Field Pinch (RFP) are all of the same order, and since these devices possess finite beta, there exists no universally small parameter in which to expand the basic equations. Instead, the full equations are integrated, using care to separate compressible and incompressible motions as much as possible[3,4]. These simulations reproduce some features of present experiments, but the next generation of computers is clearly needed here.

To make these three-dimensional codes applicable to more general geometries (e.g. stellarators) and to simultaneously include enough effects to ensure a complete description of the important physics effects (e.g. parallel heat transport, compressibility, finite larmor radius effects, and smaller values of resistivity) requires a machine with about 10 times the CPU speed of the Cray 1 as well as a large memory, e.g. the Cray 2.

Particle and Hybrid Codes

In many cases fluid models are not adequate to describe plasma behavior, for it is necessary to consider microscopic effects, i.e., the effects of the way particles are distributed in velocity. Numerically this is most often accomplished through particle codes[5-7]. Fully nonlinear kinetic ion and electron simulations in 2-D Cartesian geometry have been carried out over the last decade. In the past, Cartesian geometry was not a major physics limitation even with the obvious cylindrical and toroidal nature of experiments, because these models necessarily dealt with length and time scales on the order of the electron gyroradius and plasma oscillation period for stability. Resolving such length and time scales meant that any realistic macroscopic dimension could be considered infinite. With the increase of grid resolution allowed by improved computers and methodology, the scope of particle simulations has grown to encompass nonlocal effects and more realistic geometries.

On the present computers, large scale particle simulations in 2-1/2D and 3D are mainly limited by the size of the maximum fast memory of the Cray 1 (of the order of 1 M words, or 2 M for the Cray 1S). Experimentally relevant physics problems in magnetic confinement have important three-dimensional aspects, such as in the multiple-helicity interaction of collisionless tearing modes and in the drift wave turbulence in sheared magnetic fields; the 64M word memory of the Cray 2 and its vector addressing will greatly enhance these simulations.

Particle-fluid hybrid models have become important in the last five years. A typical hybrid model represents the ion components as kinetic species and the electrons as a fluid in order to eliminate some or all fast electron

frequencies and short length scales. Recent progress with hybrid models is impressive but is still quite computationally expensive (typically taking roughly two to four times more Cray CPU time than does an MHD code of equal dimensionality).

Fokker-Planck Codes

In the simulation of magnetically confined plasmas where the ions are not Maxwellian and where a knowledge of the distribution functions is important, kinetic equations must be solved. At number densities and energies typical of mirror machines, end losses are due primarily to the scattering of charged particles into the loss cones in velocity space by classical Coulomb collisions. The kinetic equation describing this process is the Boltzmann equation with Fokker-Planck collision terms. The heating of and current generation in plasmas by energetic neutral beams and microwaves, the thermalization of alpha particles in DT plasmas, the study of runaway electrons and ions in tokamaks, and the performance of two-energy component fusion reactors are other examples where the solution of the Fokker-Planck equation is required.

The problem is to solve a nonlinear, time-dependent partial differential equation for the distribution function of each charged species in the plasma, as functions of six phase space variables (three spatial coordinates and three velocity coordinates). Such an equation, even for a single species, exceeds the capability of any present computer, so several simplifying assumptions are required to treat the problem.

With the advent of much more powerful neutral beams, it is now possible to consider neutral-beam-driven tokamak fusion reactors[8]. For such devices, three operating regimes can be considered: (1) the beam-driven thermonuclear reactor, (2) the two-energy component torus (TCT), and (3) the energetic-ion-reactor, e.g., the counterstreaming ion torus (CIT). In order to study reactors in regimes (2) or (3), a non-linear Fokker-Planck model must be used because most of the fusion energy is produced by beam-beam or beam-plasma reactions. Furthermore, when co and counter injection are used, or major

radius compression is employed, a two velocity-space dimensional Fokker-Planck operator is required[9,10].

An example of an important 3-D (r,v,θ) calculation which is beyond the capabilities of the Cray 1 is the modeling of the transport of electron energy out of a tokamak due to the combined effects of a stochastic magnetic field and a radial ambipolar field coupled to Coulomb collisions. This problem is both nonlinear and essentially 3-D. Using an implicit scheme employing a 3-D ICCG matrix inversion package, assuming a mesh of about 120,000 points (a minimum for a physically reasonable 3-D calculation), and a cost of 1.5×10^{-3} seconds per time step per mesh point on the Cray 1, and assuming that a calculation requires 200 time steps, the amount of Cray 1 computer time required is about 10 hours, generally an unacceptable amount of time for a single run. Incidentally, the total of storage required would be about 50% greater than the matrix size or about 3.4×10^6 words. This could be accommodated on the Cray 2.

SUMMARY

In summary, as the fusion program has advanced rapidly in the last few years with the development of more sophisticated theory and experiment, computational requirements for accuracy and realism have increased to the point that Cray 2 capabilities and beyond are required. New features of the machines will allow vectorization of Monte-Carlo, finite element codes, and others which have been scalar until now, a gain of 10 in speed. When programmed to also take advantage of multiprocessing, they will be another factor of 10 faster. This will make revolutionary changes in the importance of such techniques.

It is not possible to define a performance level that represents the ultimate capability for fusion studies. Each successive generation of supercomputers has been exploited to produce more realistic results. Codes to exploit the new hardware capabilities are typically under development before the hardware is actually installed. It is safe to assert that the fusion computing community can effectively use the best performance that the supercomputer manufacturer's are capable of providing for the forseeable future.

ENERGY RESEARCH

INTRODUCTION

During FY84, the Department of Energy (DOE) created the Energy Sciences Advanced Computation activity and established, as its major program, a supercomputer access program. This program was initiated as the result of various panels which had investigated the availability of modern supercomputer resources to the scientific research community within the U.S. and to the DOE research community in particular. It was found that the current availability of modern supercomputer resources within the U.S. fell far short of the amount of these resources needed by the research community and it was also found that modern supercomputers themselves do not have sufficient capability to address many of the computational needs of this community. During FY84 a requirement analysis was conducted throughout the research community which is funded by the Office of Energy Research (ER), and this analysis verified that several Class VI computer systems would be needed to begin satisfying this supressed demand[11].

The disciplines with supercomputing needs include High Energy Physics, Nuclear Physics, Chemical and Materials Sciences, Engineering and Applied Mathemataical Sciences, Geological and Meteorological Sciences and the Biological and related sciences. Extensive computing requirements in these fields have been identified, however, new problem areas are continually being uncovered and the magnitude of the latest demand for supercomputing in the ER programs is just beginning to be understood.

The purpose of the Energy Sciences Advanced Computation Supercomputer Access Program is to provide nationwide high-speed network access to modern centralized facilities within the constraints of budgetary resources. In order to begin addressing this access problem as quickly and as economically as possible, ER decided to utilize the existing National Magnetic Fusion Energy Computer Center (NMFECC) and its installed high-speed satellite network, described earlier, across all ER programs. Because the NMFECC satellite network was already accessible at many DOE laboratories and universities and because this network provides gateways to other networks, such as ARPANET and

TYMNET, many researchers were able to gain access to the NMFECC facilities with very little lead time and minimal additional cost.

For fiscal year 1985, the Office of Energy Research is funding the Cray X-MP/2 computer system installed at the NMFECC in November 1984, to further expand the availability of supercomputer resources to the non-fusion ER programs. This system addresses the near term capability and capacity needs. The Office of Energy Research is requesting funds to replace this Cray X-MP/2 system with a more advanced Class VII system in FY87 in order to provide the capabilities required[11]. The Class VII system will be acquired through a competitive procurement at a time when U.S. vendors are expected to market at least three systems of this capability.

THE NEED FOR MORE POWERFUL COMPUTERS

Historically, scientists who use supercomputers have constrained their numerical simulations to an average execution time of about ten hours. This constraint reflects the scientist's need to make daily progress. Thus, the amount of complexity incorporated in models is scaled to the computer's ability to produce results in about a ten-hour execution time. The capability of a supercomputer dictates the amount of complexity that can be treated. Because of this limitataion, scientists engaged in large-scale numerical simulation have continually sough bigger and faster computers. Today, scientists engaged in energy research need supercomputers that are up to 200 times faster than state-of-the-art equipment.

In order to understand the requirements for more powerful computers, we must explore the generic reasons for having increased computational speed and storage.

Dimensionality. The real world exists in three space dimensions plus time. If computational models reflected the real world exactly and completely, they would treat all four of these dimensions and other parameters that are equivalent to additional dimensions. With current computers, it is possible to treat two space dimensions and time for some problem types, three space dimensions for others, and three space dimensions plus time for a very limited

set of problems. Speed increases of about a factor of 200 in this decade are needed to allow researchers to solve urgent multidimensional problems that are now intractable.

Resolution. Every region of space contains infinitely many points. Thus, the first step in modeling any natural phenomenon is to approximate the space with a finite set of zones, each of which requires a number of calculations. Increasing the number of zones means we can determine more completely and accurately what is happening in any environment, but the computational time grows very rapidly. For example, in a two-dimensional time-dependent model, the running time grows in proportional to the third power of the increase in resolution; increasing the number of zones by just a factor of 2 would increase the time to complete the problem by a factor of 8. Many complex problems now run up to 100 hours, so it is clear that resolution increases of even relatively small factors can overwhelm the capabilities of current supercomputers.

Physics. All computational models dealing with the frontiers of science and technology make simplfying assumptions about the laws of physics in order to keep the calculations from running too long. In some models, including just one additional physical effect can increase running time by a factor of 10. Faster supercomputers with much larger memories will permit researchers to solve problems that cannot now be economically solved.

Combination effects. Although dimensionality, resolution, and physics each have powerful effects on running time by themselves, the overall needs are derived from combinations of these effects. The highly complex problems now being studied in energy research programs, require computational models with higher dimensionality, and with higher resolution, and with more physics.

HIGH ENERGY AND NUCLEAR PHYSICS

The requirement for computers capable of meeting the data reduction needs of a high energy physics laboratory has, historically, been so great that all other computing requirements could be met without significantly impacting the large central facility. However, in the decade of the '80's, several new

computational needs have appeared which require the unique capabilities of supercomputers and clearly require capabilities presently associated with Class VII systems.

The theoretical high energy physics community represents an important class of users with very large computational needs. This is primarily due to the rapid rise of computational quantum field theory, particularly in numerical studies of lattice gauge theory. To put this development in perspective, it should be noted that computer simulation is a generic numerical tool for studying the behavior of particles and fields, and its importance does not rest on any particular fashion nor on the currency of any particular theoretical idea. The ability to carry out such calculations is primarily a result of the rapid increase in available computer power, and as such, it represents a permanent change in the way theoretical physics is done. The needs here fall into two distinct categories. The first category includes the more traditional forms of theoretical computation such as numerical integration, solution of integral or differential equations, calculation of Feynman diagrams, etc. The second category is the large scale numerical simulation of quantum field theory on a lattice. These calculations are highly CPU intensive. The lattice gauge theory algorithms are relatively simple, repetitive and easily vectorizable. Thus, they are well suited to a variety of parallel and pipelined architectures provided that a large, faster accessed memory is available. Even low statistics calculations on modest sized lattices require the equivalent of tens of CRAY hours.

Two newly emerging needs for computer power beyond the scope of Class VI systems are from the accelerator and experiment design communities. An example of an accelerator design requirement is for the turn-by-turn simulation of potential designs for the new superconducting super collider (SSC) accelerator currently in conceptual design. The integrated time needs here are CPU times measured in CRAY-1 equivalent years.

An example of the experiment-design-related requirement is the full simulation of Monte Carlo events in a colliding beam detector system. The number of simulated events run should, ideally, be substantially greater than the number of real physics events to be analyzed. Furthermore, since

experimental results may change the way a detector is tuned, it may be necessary to make the simulations concurrently with the taking of data, i.e., when the data reduction computers are most fully loaded.

Experimental high energy physics data reduction, which has heretofore used standard general purpose computers, also needs a new genertion of computers. The generation of detectors now just coming into use necessarily gather data at very high rates in order to extract the physics of interest from the enormously large accompanying backgrounds. The volume of data collected from these new experiments is several orders of magnitude larger than in experiments performed in the 1980 period. The computational problems are enormous and new classes of supercomputers along with special purpose processors appear to be the only practical way in which to satisfy these unfilled computational needs.

Monte Carlo simulations of lattice gauge theories, and more specifically of Quantum Chromo-dynamics (QCD), while not being the only calculations of interest of particle theory, are presently the most demanding in computational resources and the most likely to produce quantitative predictions. Two essential elements enter into these calculations:

a) the generation of gauge field configurations distributed according to the $\exp\{-S\}$ measure;
b) the evaluation of quark propagators in the background of the gauge fields provided by the above configurations.

The degrees of freedom are made discrete by introduction of a (usually) hypercubical lattice. A lattice extending for n_s sites in the spatial directions and n_t sites in the temporal one entails $4n_s^3n_t$ gauge dynamical variables associated with the links of the lattice. At present, we lack a quantitative understanding of any collective excitation which may dominate the functional integrals. Therefore all of the above link variables must be treated on the same footing. The lattice must be sufficiently big to contain a hadron, and provide enough resolution so that a lattice with the same physical volume but a finer subdivision would lead to essentially unmodified results (notion of scaling toward the continuum limit). Let us assume, to fix ideas, that the lattice extends for 10 sites in all spatial directions and 20 in the

temporal one. This give a total of 80,000 link variables, i.e., 80,000 SU(3) matrices which must be kept in the memory of the computer for a simulation of QCD. Thus a lattice configuration corresponds to 80,000 x 18 real numbers = 1,440,000 words of memory. To proceed from one configuration to the next all 80,000 link gauge variables must be "upgraded". The upgrading of a single variable involves on the order of 4000 arithmetic operations. We thus obtain an operation count of $\approx$ 320 million to generate a new configuration. Typically hundreds or thousands of configurations must be generated to produce meaningful results. The calculations of the quark propagators are about as demanding in computational resources.

In conclusion a computing center which would serve the interest of high energy theorists should be endowed with one of the most powerful mainframes available, both in computational speed and in memory size, such as the ER Class VII system proposed for FY87.

BASIC ENERGY SCIENCE

Material Sciences

The development and proliferation of investigations of diverse material systems and phenomena via computer simulation and modeling is a rich field of scientific endeavour anchored in the physical sciences (with cross-fertilization links to advances in applied mathematics and computer science), made possible singularly by the advent of high-powered computers. Computer simulations provide information about phenomena and processes in material systems with refined microscopic spatial and temporal resolution and enable investigations of the dynamical evolution of complex systems under extreme conditions where data from experiments or other methods of investigation is not attainable. In addition such studies provide benchmarks for critical testing and refinement of theoretical concepts and aid in the interpretation of experimental observations.

Current simulation methods involve the generation and analysis of phase-space trajectories of an interacting many-particle system either by the direct numerical integration of the equations of motion (molecular dynamics-MD,

and reaction-trajectory-TJ, methods) or via the sampling of phase-space configurations (Monte Carlo-MC). In either case the many-body nature of the systems under study and the statistical modes of analyses dictate the necessity for extended computer time and storage capabilities.

The wide range of materials system investigated by computer simulations include: the equilibrim and non-equilibrium structure and dynamics of materials at different states of aggregation (solids and liquids) and the kinetics and dynamics of phase transformations; properties of metastable systems (supercooled liquids, quenched liquids, gasses); homo and multicomponent materials; ordered versus disordered (amorphous) solids; surfaces; interfaces and inter-phase interfaces, i.e, solid-solid (superlattices and coherent structures), solid-liquid (epitaxial crystal growth and homogeneous nucleation), solid-gas (molecular beam epitaxy, heterogeneous catalysis).

Simulation studies on these systems allow investigation of structural and dynamical characteristics, kinetics and dynamics of phase-transformations, transport and non-linear phenomena (heat, matter, electrical), diffusion processes and reaction dynamics. Furthermore modifications of the intrinsic properties of condensed matter systems and phenomena (such as fracture, solid transformations, plastic flow), due to external fields (mechanical stress, heat gradient etc.) can be investigated. In addition to an improved understanding of existing material systems, simulation studies could serve as the impetus for exploration of methods of preparation and growth of novel materials.

Underlying simulation studies of extended condensed matter systems is the notion that the properties of the "calculational sample" on which the simulation is carried out, extended via the commonly used periodic boundary conditions, are a faithful representation of the nature of the macroscopic system. Among the factors which dictate the size of the "calculational sample" are the ranges of interparticle interaction potentials and fluctuation wavelengths. Thus for example the MD simulation of the structural and dynamical properties of a solid simple metal (e.g., Al) requires a system containing ~2000 particles; the simulation of binary liquid metals and supercooled liquids require an even larger sample due to concentration fluctuations. Simulations of stressed crystals, fracture and plastic flow,

shock wave propagation, the dynamics of melting and hydrodynamical phenomena would require systems where the number of particles would be 5000-10,000. It should be noted that in the presence of long range and realistic multibody forces the computing time grows as a (low) power of the number of particles. Such requirements necessitate memory capacity beyond CRAY-1S capability and large increases in computational speed.

A critical input in materials simulations is the interparticle interaction potential. A faithful simulation requires the calculation of such potentials via pseudo-potential methods which, for metalic systems, depends upon the thermodynamic state variables (density, temperature, pressure). Simulations of nonequilibrium phenomena (such as solidification, quenching etc.) in which the state variables themselves evolve in time require a self-consistent adjustment of the interaction potentials along with the dynamical evolution of the system.

The coupled complexities of size and interaction potential calculations make such simulations prohibitive on the Cray 1. Furthermore, the magnitude of such simulations dictate substantial time requirements, for example, 80 minutes of CRAY-1 time allow the generation of ~5000 integration time steps for a system containing 1500 particles interacting via simple truncated Lennard-Jones potentials, with a fully optimized code. Note that this is the least demanding model from a computational point of view. A typical study of the solidification of such a system requires 50,000 integration time steps. It should be emphasized that the above considerations are dictated by the nature of the physical systems and phenomena and cannot be compromised by approximate treatments which will prejudice and distort the simulation results. Thus progress in this field can be made with substantial access to the Class VII computing facilities proposed for FY87.

REFERENCES

1. B.I. Cohen and J. Killeen, Phys. Today (May 1983).

2. M.N. Rosenbluth, D.A. Monticello, H.R. Strauss, R.D. White, Phys. Fluids 19, p. 1987 (1976).

3. D. Schnack and J. Killeen, Nucl. Fusion 19, p. 877 (1979).

4. E.J. Caramana, R.A. Nebel, D. Schnack, Phys. Fluids 25, p. 1305 (1983).

5. J. Dawson, H. Okuda, and B. Rosen, Methods in Comp. Phys. 16, p. 282 (1976).

6. A.B. Langdon and B. Lasinski, Methods in Comp. Phys. 16, p. 327 (1976).

7. C. Nielson and H. Lewis, Methods in Comp. Phys. 16, p. 367 (1976).

8. D.L. Jassby, Nucl. Fusion 17, p. 309 (1977).

9. J. Killeen, A.A. Mirin, and M.E. Rensink, Methods in Comp. Phys. 16, p. 389 (1976).

10. J. Killeen, A.A. Mirin, and M.G. McCoy, Modern Plasma Physics, IAEA-SMR-61/110, p. 395 (1981).

11. "The Role of Supercomputers in Energy Research Programs," U.S. Department of Energy, DOE/ER-0218, February 1985.

Mr. FUQUA. Thank you very much.

Dr. Johnson.

[The biographical sketch of Dr. Johnson follows:]

ROBERT MERRILL JOHNSON

PERSONAL DATA

Born September 17, 1926, Detroit, Michigan. Married, four children. Home Address: 306 Saratoga Drive, Tallahassee, Florida 32312 (Phone: 904-385-4164). Business Address: Graduate Studies & Research, 408 Westcott Building, Florida State University, Tallahassee, Florida 32306 (Phone: 904-644-3500). Present Position: Dean, Graduate Studies & Research, Florida State University.

EDUCATION

Ph.D., Michigan State University, 1957, major in physiology, minor in biochemistry; M.S., University of Detroit, 1953, major in biology; B.S., University of Detroit, 1951, major in biology.

EMPLOYMENT HISTORY

Military Service: Army Air Force; June 1944 through October 1946 (one year of Air Force training was spent at Michigan College of Mining and Technology in the ASTP Program).

The Army Exchange Service: Munich, Germany; November 1946 through June 1947.

Manufacturer's Agent: Baker and Collinson, 1200 Mt. Elliott, Detroit, Michigan, January 1953 through August 1954.

University of Detroit: 1947-53; Graduate Teaching Fellow, Department of Biology, 1951-52.

Michigan State University: 1954-57; Graduate Assistant, Department of Physiology, 1954-57.

Colorado State University, 1957-66: Instructor, Assistant Professor, Associate Professor, Full Professor, Department of Physiology, 1957-62; Assistant Director, Colorado State University Research Foundation, 1959-62; Director, Facilities Development; Professor, Department of Physiology, 1964-66.

National Science Foundation: Associate Program Director, Specialized Facilities and Special Programs, Division of Biological and Medical Sciences, July 1962 through October 1962 (on special leave from Colorado State University).

Program Director, Division of Institutional Programs, November 1962 through June 1964 (on special leave from Colorado State University).

Staff Associate, Institutional Relations, September 1966 through September 1968.

Florida State University: Dean, Graduate Studies and Research, and Professor of Biological Science, September 1968 to present.

HONORARY AND PROFESSIONAL ORGANIZATIONS (CURRENT)

American Physiological Society; American Association for the Advancement of Science (Fellow); Electron Microscope Society of America; Sigma Chi Iota; *Who's Who in America,* 39th Edition, 1976-77, Volume 1; 1977-78; 1978-79; 1980-81; 1981-82; 1983-84; *Who's Who in the World,* 3rd. Edition, 1976-77; 4th Edition, 1978-79.

STATEMENT OF DR. ROBERT M. JOHNSON, DEAN, GRADUATE STUDIES AND RESEARCH, FLORIDA STATE UNIVERSITY, TALLAHASSEE, FL

Dr. JOHNSON. Thank you.

I am very pleased to be here speaking before the committee. I am not quite sure why I am here, because I am not a director of one of these centers. The director of our center is Dr. Joseph Lannutti, but I will do my best. And if I run into technical problems, I think Joe is here and a few of the other people who can help me out.

Before I proceed, I would certainly like to take this opportunity to thank Congressman Fuqua and, of course, the members of his

committee and Congress, and our partners in the State of Florida, and DOE, for helping us develop this center here at Florida State University.

In my prepared statement, I think I gave you a copy of a manuscript that we have called "Research in Review." I hope that can be incorporated into the record. I think that speaks very well as to where we are in the supercomputer research today—headed by Dr. Joseph Lannutti.

The staff, as you can see on page 5 of that, is building up very rapidly. There are eight groups of off-campus scientists engaged in energy research projects for the Department of Energy, along with five groups of FSU researchers in the energy research area. So we are moving very, very rapidly in building up that whole center.

As far as the center is concerned, it is pretty well documented in this publication that we put out, prepared by our research editor, Frank Stephenson, so I will not dwell on that issue much longer.

Mr. FUQUA. Do you have—I do not think everyone has a copy of that.

Dr. JOHNSON. I have some extras. I anticipated that you did not all have copies.

Mr. Chairman, I would like now to turn to a subject which has been, I guess, pretty well worked over here at this conference today already, this hearing, and that is the serious problem facing us in making these resources available to other members of the university community; that problem is the lack of access by the universities to sufficient high-bandwidth communication capability. If we assume that all the good researchers are found only at major universities, we do not only do a disservice to the smaller institutions which have very excellent personnel, but we begin to ensure that eventually such a statement becomes true, to the detriment of our educational system. Good scientists with a need for supercomputer access are found at many places, and some of these places may not be large enough to support a dedicated supercomputer. Yet, what a disservice we would render this Nation to suggest for a moment that they should not have supercomputer access.

At a previous hearing of this committee, the suggestion was made that every campus in the Nation should have a supercomputer, and perhaps some day the technology will evolve to that point where it could be economically feasible, but it is not feasible today. But communications capability to make the existing centers available to other universities is feasible now. Low-cost earth stations consisting of an up-link and a down-link are available for a few thousand dollars, and certainly we have unparalleled skill at placing communications satellites in the heavens. I believe it is a legitimate function of this committee to examine whether or not this existing technology can effectively be utilized to make supercomputers accessible to qualified scientists and researchers at all institutions, and not just the select few.

The primary problem in supercomputer software development today is a methodology for effectively utilizing multiple processors. When this problem is solved, there will be a quantum leap in effective use of supercomputers.

Mr. Chairman, I ask that you consider for a moment what it would mean to have available a 50- to 60-megabit communications

path linking universities. Such a link would not only have the capability of carrying large data sets for supercomputer use, but would also serve to carry telephone conversations, television signals, electronic mail, data bases for small computers, and in general to serve as a pipe for linking these centers of learning.

It is well accepted, I believe, that such interaction between our scientists and researchers provides synergistic benefits far beyond what might be expected. I refer in particular to the aspect of training. I believe that having a suitable communications link is an absolutely vital aspect of providing training in the use of these machines. It is technologically feasible, I believe, for smaller colleges to make use of the facilities at other institutions by proper communications links.

In this time of economic belt-tightening, we must look to the most efficient method of utilizing our resources. Supercomputers are but one example of resources which can be profitably shared. I believe that other examples can be cited in laboratory instrumentation, in medical tools, and in libraries, to name but a few.

Such a concept of scattered resources which can be shared leads inevitably to a communications access medium, such as I have attempted, with my limited technical knowledge, to describe. However, lest I be thought completely uninformed in such a highly technical area, let me hasten to assure you that I am speaking in terms of technology which already exists today, in the form of Time Division Multiple Access [TDMA] which allows many Earth stations to share a single satellite transponder.

It is my understanding that several TDMA systems have already been put into operational service, and more are planned. Unfortunately, however, these are not generally available to universities, and it is not feasible for a single university to bear the cost of such a system. With the leadership of this committee in the concept of shared Government-industry efforts, I think that such an endeavor is a proper effort for Government to endorse and for industry to support.

We, at Florida State University, have been working with our scientific colleagues within SURA—that is the Southeastern Universities Research Association—to put together such a plan for over a year, with the idea that SURA would be an ideal vehicle for a pilot project to demonstrate the feasibility of such an endeavor. However, unfortunately we were not able to get the commitment from private industry which we needed to carry that particular plan forward, although SURA does have a less ambitious plan for a communications network called SURANET before the NSF now, which we hope will be funded. I am convinced that unless Government is willing to take the lead, no one university or group of universities is going to be able to shoulder that burden.

The other subject which we have touched on here in the hearing today is on training, and this will take just one short sentence or two on that. I believe that the current shortage of skilled supercomputer professionals is not in the best interest of our Nation. I think the efforts the National Science Foundation has taken to develop their workshops and summer workshops, the three that we heard that were established this summer, needs to be expanded. And I would certainly like to see that expanded into other depart-

ments, such as our own Department of Energy, helping to fund our program, and other defense agencies, NASA, and so forth.

The future of our young people in this technology is something which we must all address. Therefore, I recommend that such an extension of these programs be considered in light of the problems facing us today, with participation by those universities which have supercomputers available and are willing to make facilities available for such a training program.

Thank you.

[The prepared statement of Dr. Johnson follows:]

Statement of Robert M. Johnson, Dean of Graduate Studies and Research, Florida State University

Mr. Chairman and members of the Committee:

My name is Robert M. Johnson, and I am Dean of Graduate Studies and Research at Florida State University. I am very pleased to have this opportunity to appear before you today to discuss what we at Florida State University see as vital long range federal policies to enhance research, training and applications of supercomputer technology.

It seems only yesterday that this Committee was breaking new ground in suggesting that U.S. scientists and researchers were being hurt badly by not having access to supercomputer technology -- a resource which is readily available to researchers in European and Japanese universities. I am pleased to report to you today that with your encouragement and leadership, we can see some light at the end of the tunnel. Florida State University has joined a growing list of universities who now have access to supercomputer technology.

Our Supercomputer Computations Research Institute, or SCRI as we refer to it, is busily accumulating the people -- the scientists, the specialists, the support staff -- to show scientists how to utilize the supercomputer in their areas. In addition, our SCRI-connected faculty members, funded by the State of Florida as a contribution to this joint effort, are beginning to serve a catalyst role in stimulating interest among FSU researchers. Supercomputer time itself is a commodity already in short supply, as eight groups of off-campus scientists engaged in

energy research projects for the Department of Energy, along with five groups of FSU researchers in the energy research area have already been assigned time by the DOE. The off-campus users represent research teams of from five to fifteen people each, and include four high-energy physicists, five nuclear physicists, three specialists in materials research, two chemists, three applied mathematicians, and a molecular biophysicist. On campus, DOE users include our high energy physics group, and Geophysical Fluid Dynamics Institute. So we are already putting this new tool to good use. You have been furnished with a copy of our publication "Research In Review," which goes into more detail on the subject, covering the research areas in more depth.

Mr. Chairman, I would like to turn to what I see as the most serious problem facing us in making those resources available to other members of the university community; that problem is the lack of access by the universities to sufficient high-bandwidth communications capability. If we assume that all the good researchers are found only at major universities, we not only do a disservice to the smaller institutions which have very excellent personnel, but we begin to ensure that eventually such a statement becomes true, to the detriment of our educational system. Good scientists with a need for supercomputer access are found at many places, and some of these places may not be large enough to support a dedicated supercomputer. Yet what a disservice we would render this nation to suggest for a moment that they should not have supercomputer access. At a previous hearing of this Committee the suggestion was made that every campus in the nation

should have a supercomputer, and perhaps some day the technology will evolve to the point where that is economically feasible. It is not feasible today, but communications capability to make the existing centers available to other institutions is feasible now. Low cost earth stations consisting of an up-link and down-link are available for a few thousand dollars, and certainly we have unparalleled skill at placing communications satellites in the heavens. I believe it is a legitimate function of this Committee to examine whether or not this existing technology can effectively be utilized to make supercomputers accessible to qualified scientists and researchers at all institutions, and not just the select few. The primary problem in supercomputer software development today is a methodology for effectively utilizing multiple processors. When this problem is solved, there will be a quantum leap in effective use of supercomputers. It may well be that a researcher at some small university somewhere will be the key to that quantum leap -- IF he or she can get access to a supercomputer. Communications will be the key to that access.

There has been a lot of effort, by a lot of people, put into establishing the methodology of user-oriented communications networks, which have been instrumental in bringing the technology to where it is today. The various networks such as ARPANET, BITNET, CSNET and MFENET have taught us much. The planning and effort that has gone into SCIENCENET has consolidated a lot of these ideas, and has brought in new ones. The time has come to adapt this accumulated knowledge to meet the high bandwidth

requirements of supercomputer technology. Mr. Chairman, I ask that you consider for a moment what it would mean to have available a 50 to 60 megabit communications path linking the universities. Such a link would not only have the capability of carrying large data sets for supercomputer use but would also serve to carry telephone conversations, television signals, electronic mail, databases for small computers, and in general to serve as a pipe for linking these centers of learning. It is well accepted, I believe, that such interaction between our scientists and researchers provides synergistic benefits far beyond what might be expected. I refer in particular to the aspect of training. I believe that having a suitable communications link is an absolutely vital aspect of providing training in the use of these machines. It is technologically feasible, I believe, for smaller colleges to make use of the facilities at other institutions by proper communications links. In this time of economic belt-tightening, we must look to the most efficient method of utilizing our resources. Supercomputers are but one example of resources which can be profitably shared. I believe that other examples can be cited in laboratory instrumentation, in medical tools, and in libraries, to name but a few. Such a concept of scattered resources which can be shared leads inevitably to a communications access medium such as I have attempted, with my limited technical knowledge, to describe. However, lest I be thought completely uninformed in such a highly technical area, let me hasten to assure you that I am speaking in terms of technology which already exists, in the form of Time Division Multiple Access which allows many earth stations to share a

single satellite transponder. It is my understanding that several TDMA systems have already been put into operational service, and more are planned. Unfortunately, however, these are not generally available to universities, and it is not feasible for a single university to bear the cost of such a system. With the leadership of this committee in the concept of shared government-industry efforts, I think that such an endeavor is a proper effort for government to endorse, and for industry to support. We at Florida State University have been working with our scientific colleagues within SURA to put together such a plan for over a year, with the idea that SURA would be an ideal vehicle for a pilot project to demonstrate the feasibility of such an endeavor. Unfortunately, we were not able to get the commitment from private industry which we needed to carry that particular plan forward, although SURA does have a less ambitious plan for a communications network called SURANET before NSF now, which we hope will be funded. I am convinced that unless government is willing to take the lead, no one university or group of universities is going to be able to shoulder the burden.

I am reminded, Mr. Chairman, of a story which I have been told, and which I believe to be accurate. There exists one institution which has a very scarce resource in the form of a scanning device which can provide data to assist in the diagnosis of medical problems. The analysis of the data, however, is very complicated, and still another institution has the leading expert in interpreting the data. The present practice is to collect the data and send it by mail to the other institution, where in due course it is analyzed, and the results returned. However, with a

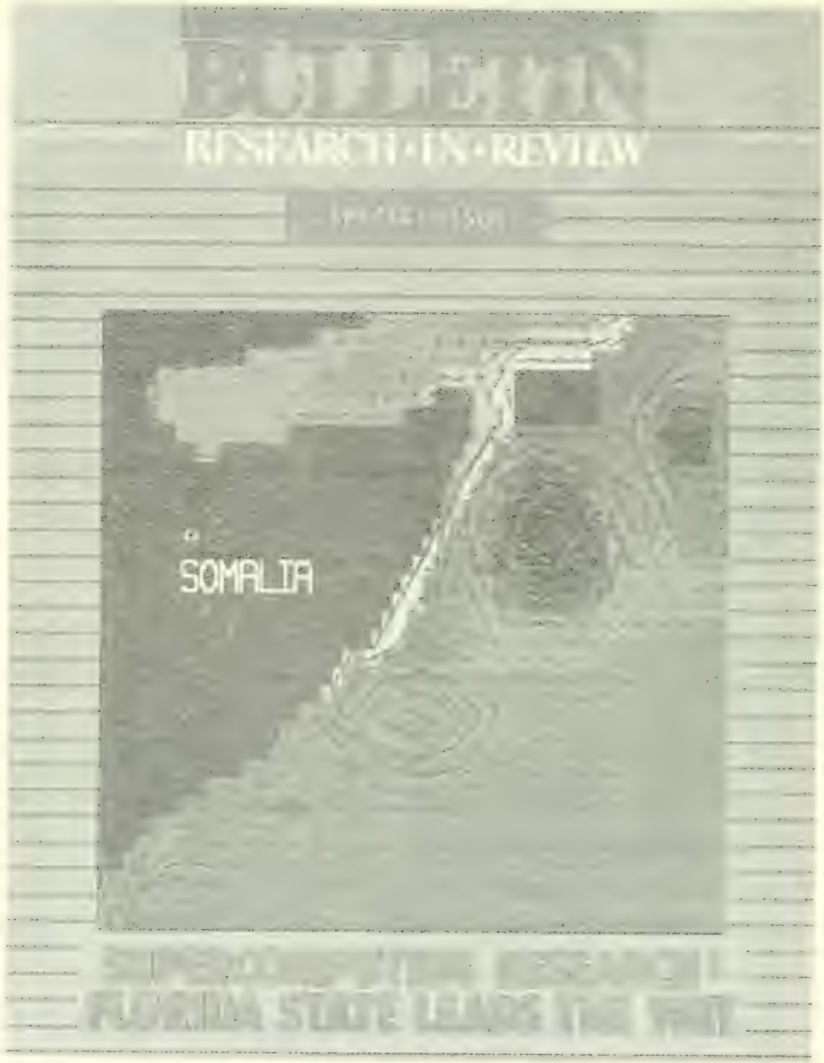
high speed communications link, it would be possible to immediately send the data, and to have the results returned by an electronic medium. Obviously one such example is not justification for installation of an expensive and intricate system such as I have described.. However, I believe that many such examples could be cited by our scientists and researchers, covering many different and varied areas of expertise. The sum of the existing needs, plus the new needs which would be generated as experience is gained in using such a system, would pay dividends far in excess of the cost of such a system in terms of knowledge, of efficiency, and in usage of shared resources.

Let me turn again to the subject of training. I believe that the current shortage of skilled supercomputer professionals is not in the best interest of our nation. I speak, of course, not of someone who merely has used a supercomputer, but of someone who knows how to take advantage of the characteristics of the machine to use it in an optimum fashion. As you know, it is the intention of Florida State University to serve as a primary training center, among the other aspects of our activity, but I think again that it is a proper function of government to assist and encourage training efforts to the maximum extent possible. I recall that in the 1960's there was a National Science Foundation sponsored summer program for bright youngsters at Florida State University which identified several of the people who subsequently went on to participate as professional employees in making our Computing Center the success it is. It is my understanding that NSF is now in the process of making such a program

available for summer institutes to use the supercomputer facilities they have sponsored. Certainly this is desirable, and I applaud their efforts. I would like, however, to see these efforts extended to other agencies, such as the Department of Energy, as well. We stand ready, also, to assist the National Science Foundation or any other agency in implementing such a training program utilizing the supercomputer facilities at Florida State University. The future of our young people in this technology is something which we must all address. Therefore, I would recommend that such an extension of such a program be considered in light of the problems facing us today, with participation by those universities which have supercomputers available and are willing to make facilities available for such a training program.

Knowing the time constraints this Committee is under, I have attempted to keep my remarks brief, although in addressing these issues I speak of things which concern me greatly, and about which I could go on at great length. But I will be brief, Mr. Chairman, and in closing let me urge that each of you consider not just my words, but the intent behind them, which is to relay to you on behalf of myself and my colleagues the deep appreciation that we have for your efforts, and to set forth to you the issues as we see them.

Thank you for your attention and your consideration of these issues. I will be happy to answer any questions, or to go into such further detail as you desire.


RESEARCH • IN • REVIEW
SOMALIA

FLORIDA·STATE·UNIVERSITY

BULLETIN

RESEARCH·IN·REVIEW

ON THE COVER

"The Great Whirl," off the coast of Somalia in the Arabian Sea, has been known to mariners since ancient times. The giant current, along with its temperature patterns, is being studied by FSU oceanographer Dr. Mark Luther using modeling techniques designed on the Cyber 205 supercomputer. In the graph, red indicates warmer water. A specialist in the numerical modeling of oceans, Luther is part of a team of FSU scientists using supercomputers to study the relationship between the earth's ocean and its atmosphere. Designing models, based on millions of real weather observations, is the only way scientists can study the nature of such immense natural systems. Supercomputers are essential tools in the work.

Florida State University Bulletin *(USPS 022-360) is published five times each year (twice in March, and in June, September and December) with editorial offices at 408 Longmire and 322 Bryan Hall, Florida State University, Tallahassee, Florida 32306. Second class postage paid at Tallahassee, Florida.* ***Postmaster:*** *Send address changes to Florida State University Bulletin, 322 Bryan Hall, Florida State University, Tallahassee, Florida 32306. All issues except the second March issue are published by the Office of Graduate Studies and Research at an annual cost of $34,092 or $0.93 per copy to report and interpret research activities at Florida State University. Any written portion of this publication and any non-copyrighted artwork may be reprinted without permission. Credit for* ***Florida State University Bulletin, Research in Review*** *is requested. Correspondence should be addressed to Frank Stephenson, University Research Editor, 322 Bryan Hall (904) 644-5254 or to Dr. Robert M. Johnson, Dean, Office of Graduate Studies and Research, 408 Westcott, Florida State University, Tallahassee, Florida 32306. Opinions expresses in this publication do not necessarily reflect those of the Florida State University faculty or administration.*

EDITOR *Frank Stephenson* / WRITER *Mary Tebo* / DESIGN *Margo Shand* / PHOTOGRAPHY *Bill Langford, FSU*

THE SUPERCOMPUTER

COMPUTATIONAL

RESEARCH INSTITUTE

A MISSION, A MACHINE

FLORIDA STATE JOINS THE SUPERCOMPUTING RACE

"We're going to make believers out of a lot of people."

As two refrigerated semis slowly pulled into view on East Dirac Drive on the clear morning of March 4, an FSU engineer punched a finger into a buddy's ribs.

"There she is, ace. She's really here. *Hot damn!*" The two men hugged each other. They had been waiting months for the moment, and the trucks just weren't moving fast enough to suit them. They itched to get started unpacking 10 tons of the fastest, most powerful computer ever to cross the state line.

Within four days, technicians had tightened the last bolt in the installation of a $12 million Cyber 205 supercomputer, the heart of a $63 million research center like no Florida university or business has ever seen. Inside a week, the machine was casually munching through a mountain of numbers fed into it by FSU scientists.

The scientists watched closely for hiccups. They found a few, diagnosed as "normal." By May 15, things were running to "specs." More smiles, more handshakes. Florida State had a computer as big as they grow.

continued on next page

"Now don't misunderstand. Call it the *hardware* heart, if anything. It's just a machine, that's all. It's the people -- the scientists, the specialists, the support staff --they're the real heart of this thing. They're the ones whose imaginations are going to make the whole thing work."

When Dr. Joe Lannutti talks, it's often in whole paragraphs at once, airborne word assaults, unedited perhaps, but unfailingly sincere, even urgent-sounding. Never one to shy away from conversation, Lannutti is talking like never before these days. He was named director of Florida State's Supercomputer Computations Research Institute (SCRI) about a year before the semis pulled onto Dirac Drive, Innovation Park, Tallahassee. It's been a talkative 12 months.

"I can't overemphasize the point. It's the high-caliber people here who are going to make this happen, not a fancy machine. True, we can't work without a tool. And we've got a good one. But that's all it is, a tool. It's up to us to use it well."

Down the hall from Lannutti's fifth-floor office in the Keen Building, home of FSU's physics department, two young physicists are engaged in a lively, if wholly unintelligible discussion. Both are given to much gesturing, European-style. One is from Germany, the other from Hungary, Lannutti says. Both are new SCRI employees, part of a staff that has grown from two to 20 in six months. It's still growing.

For the last few months, Lannutti has spent much of his time sifting through dozens of applications from specialists in computer hardware, software, and communications, from researchers of varied stripe, from scientists eager to hire on at a university equipped with a first-class supercomputer. He's authorized to spend up to $1 million this year in state funds to hire 12 of the best supercomputer-oriented scientists he can find.

The work that required three months can now be done in less than a week.

In addition, thanks to industry's role in the tripartite coalition (it's the nation's first government-university-industry venture in supercomputing) he's planning to beef up an already respectable 11-person technical and research staff. By next spring, he expects SCRI (which is being heard as "scree") to have about 50 employees, not counting the dozen or so new faculty members that will be scattered around campus, doing their "salting" work in various departments.

A primary role of the SCRI-connected faculty members, Lannutti explained, will be to serve as catalysts for stimulating interest among FSU researchers in supercomputing. When it comes to applying state-of-the-art computing to their research, FSU scientists are in the same boat with all American university scientists -- only a fraction can claim first-hand experience in using supercomputers. Lannutti says that SCRI is out to change all that.

"We're going to make believers out of a lot of people. We've designed the institute to help show scientists how to make headway faster in their fields. And maybe in ways they've never dreamed of."

53-831 O -86 - 6

DR. JOSEPH LANNUTTI

Arvil Williams with the FSU Computing Center inspects the inside of FSU's newly installed Cyber 205 supercomputer.

That's just a matter of time, Lannutti believes. But for now, time — at least on the Cyber 205 — is a commodity already in short supply.

Supercomputer-starved scientists from around the country have heard the news: there's low-priced supercomputer-time available in Tallahassee. For scientists obliged to pay upwards of $2,000 an hour for domestic supercomputing time (if they can qualify for it at any price), or worse, forced to travel overseas for it, the news couldn't have been sweeter. That's what federal backers of the SCRI were counting on all along (see page 12).

Since the U.S. Department of Energy's Office of Energy Research (ER) is paying 70 percent, or about $44 million, of the institute's operating costs through 1989, the SCRI will focus on research in which ER has a direct interest. Sixty-five percent of the 205 machine's calendar year (about 6,000 hours) is to be made available to scientists with an approved ER contract (so-called ER "users"). And there are hundreds nationwide (and some abroad), all engaged in unclassified research in fields ranging from elementary particle physics to pollution control.

But if you're an ER user just getting the word, you'll have to take a number to work at FSU. Weeks before the 205 was uncrated at Innovation Park, scientists from Massachusetts to California had gobbled up most of the ER-allotted time on the machine. This year's 65 percent is now locked up by 18 ER users outside Florida and five groups of campus users, says Dr. Dennis Duke, Florida State physicist and SCRI official. "We're booked solid. And that's 24 hours a day, seven days a week," Duke said.

Each of the out-of-state names approved for time on the machine represent research teams of five to 15 people each, Duke said. The list includes four high-energy physicists, five nuclear physicists, three specialists in materials research, two chemists, three applied mathematicians and a molecular biophysicist.

FSU will use the machine to study the atmosphere and El Nino.

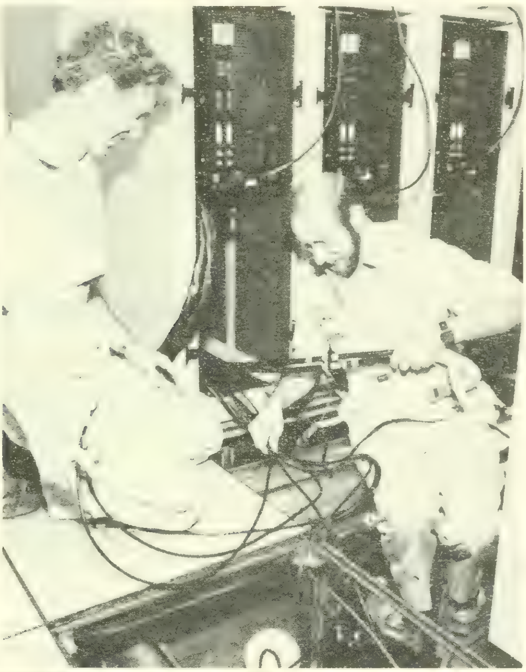
The latter, Dr. Suse Broyde from New York University, uses supercomputers to study how various cancer-causing agents attach themselves to human DNA. She told *Research in Review* that supercomputers had "opened up whole new worlds of possibilities" for her research. "The work we used to do in about three months we can now do in less than a week," she said. "We used to have so many constraints on the nature of the questions we could ask. No more." Broyde plans to use her 205 time at FSU to improve the complex computing codes required to model molecular structure.

With 400 hours of time approved for his research team's use for the next six months, Duke ranks far and away the campus' No. 1 user of the new machine. His team of three Ph.D. research scientists and three graduate students are going to use the 205 to solve immensely complicated equations in lattice gauge theory. In April, more than 80 physicists and mathematicians from around the world converged on FSU's Florida Conference Center for

DR. JIM O'BRIEN

DR. MARK LUTHER

Workmen install the Cyber 205 at Innovation Park.



an institute-sponsored conference on the subject. Duke said the theory, which physicists use to help describe the nature of subatomic particles, commonly requires the tackling of millions of numbers even in its simplest application. "This is the kind of (research) that is totally impossible to do any other way," he said. "Without a supercomputer you don't even start."

Three FSU mathematicians, Drs. John Bryant, De Witt Sumners and R.C. Lacher, in collaboration with FSU chemist Dr. Leo Mandelkern, will use 205-generated mathematical models to study how certain kinds of polymers are constructed. And an FSU newcomer, meteorologist Dr. Eric A. Smith, plans to use the 205 to continue his work in satellite meteorology, his specialty. Smith, who came to FSU from Colorado State University in April, is the first of the dozen computer-oriented scientists expected to join SCRI this year.

The remaining 35 percent of the Cyber 205's time each year (about 2,000 hours) is open to any FSU faculty member -- or any Florida university researcher for that matter -- and any outside agency whose project is approved by a campus-based steering committee. Headed by Dr. Jerry Stephens, FSU meteorologist and SCRI administrator, the nine-member committee evaluates all proposals which fall "outside the DOE fence." Among the most computer-intensive research areas in this category are the closely allied fields of oceanography and meteorology.

It's no surprise to SCRI officials then, that slated to be the heaviest "non-ER users" of the 205 are research teams led by Florida State's Dr. Jim O'Brien and Dr. T.N. Krishnamurti. Both men are internationally-known pioneers in the mathematical modeling of the world's oceans and atmosphere.

They figure prominently in a small, elite group of about 200 scientists worldwide who are proficient in designing extremely complex computer-based models that simulate real world weather and ocean behavior. Such models, which forecasters the world over depend on, require the rapid assimilation of millions of measurements taken daily by weather satellites, airborne monitoring devices and ships scattered around the globe.

O'Brien is perhaps best known for his work describing the behavior of the notorious tropical weather phenomenon called El Nino, almost a household term since the harsh winter of 1983-84. In fact, O'Brien chairs a newly formed committee called El Nino Southern Oscillation (ENSO) which, in effect, is an El Nino watch group, analogous to the hurricane watch group stationed at the National Hurricane Center in Miami.

On campus, he heads a 22-member research team called the Mesoscale Air-Sea Interaction Group, funded largely by the Office of Naval Research, NASA, and the National Science Foundation. The group is concentrating on two projects: modeling the behavior of El Nino in the mid-latitudes (what O'Brien calls "The California El Nino Problem. *It plays hell with fisheries from Baja to Alaska, among other things.*") and the behavior of the Indian Ocean. The team plans to use the Cyber 205 to improve their models and to build new ones.

One of O'Brien's key research people is Dr. Mark Luther, an oceanographer by scientific training and a supercomputer programmer by, as he puts it, "have-to." At 31, Luther typifies a new, though still rare, breed of young scientist showing up in first-rate campus research labs these days. Not only is he a classically trained scientist (Ph.D., North Carolina-Chapel Hill), he's also a highly skilled supercomputer user. He learned science at UNC, and supercomputers in a government-run lab, the National Center for Atmospheric Research (NCAR) in Boulder, Colorado. Supercomputers are just tools of his trade, Luther says.

"In large-scale work like we do, you have to learn (supercomputer) programming to do the job well," he said. "I'm really not a programmer. I do it because I have to. In this field, you learn early that supercomputers are the only way to get where you want to go." Luther's own mathematical model of the Arabian Sea, built on another Cyber 205 last year, was used as a principle test of FSU's machine during a month-long trial period.

O'Brien is excited by what ready access to first-class computing poses for his group's research. For years, supercomputing's stiff cost, plus the difficulty in finding it, have stymied both his productivity and his imagination, he said. "The 205 here means that in one instant I can catch up with everybody else in the country," he said. "Now I can start competing with the national (government) labs. I anticipate more respect for our research because now we'll be able to handle more meaningful, more substantial problems than ever before."

A five-minute walk from O'Brien's office can put you knee-deep in computer data, much of it the same sort of stuff O'Brien's group sorts through each day. It's the suite of labs consigned to T.N. Krishnamurti and his multi-national team of grad students. It's where "Dr. Krish," as he's affectionately called, routinely spends 12-hour days poring through world weather maps, computer tapes, print-outs and the like.

In May, Krish was named Florida State's Robert O. Lawton Distinguished Professor, the top tribute FSU faculty members can pay to one of their own. The honor came on the heels of another; it, too, superlative. In a January ceremony in Los Angeles, Krish was handed the American Meteorological Society's highest award, the Carl-Gustaf Rossby research medal. He was being honored for his "leadership in the analysis of global atmospheric observations."

Krish credits much of his success to supercomputers. For 15 years, he's fought for every bit of supercomputing he could lay his hands on. When he couldn't find enough in this country, which was all too often, he found it in Europe. For the past several years, his annual supercomputing bill from NCAR has averaged in the neighborhood of $500,000. In 1983, the NSF handed Krish a Creativity Award (the agency's first) and $400,000 to do with as he saw fit. He promptly bought more computer time.

The upshot is that when weather scientists talk about experts in tropical weather modeling, they talk about Krish or they don't talk at all. Dr. William Bonner, head of the National Weather Service and Krish's former classmate at the University of Chicago, says of him: "Krish has contributed more to numerical weather prediction and to understanding what goes on in the tropics more than anyone I know. One of the noteworthy things about his research is not only its all-around excellence but the fact that it deals with real problems in the atmosphere that have to be dealt with to make better forecasts."

Soon the biggest growth in applications will be in industry.

Literally hundreds of miles of computer tape, bearing weather observations collected by ships, balloons, anchored or drifting weather instruments, commercial and private aircraft -- plus satellite scans of 17-mile-thick layers of atmosphere -- pass through Krish's team's hands each year. Once digested by supercomputers, the millions of numbers yield elaborate mathematical frameworks, or models, which can be used to study the myriad phenomena that contribute to making the world's weather.

Chairman of a group called the U.S.-India Initiative in Monsoon Forecasting, Krish uses his models to help predict the seasonal fluctuations that can so drastically effect life in

DR. ERIC A. SMITH

DR. DENNIS DUKE

his native country. "My students are extremely excited by the 205 here," he said. "It will enable us to do problems we never imagined we could do. I see a whole new area of possibilities opening up for us."

They'll be very high-priced commodities when they graduate. You can bet on it."

Dennis Duke is talking about graduate students and their futures. He's talking about *his* grad students, but he really means any grad student lucky enough to get hands-on experience with a supercomputer.

"And believe me, they know it. We've had several calls from around the country already. They just phone us up, asking about applying for grad school here. Totally unsolicited. The news is getting out."

Nothing has so fired up today's crop of graduate students in physics, meteorology, engineering, computing science, math, biochemistry, oceanography and geology as have the prospects of getting an education that includes first-hand knowledge of supercomputer applications in their chosen fields. Thanks to initiatives such as SCRI, the nation's first government-industry-university partnership in supercomputing, a substantial number of budding scientists and engineers can now realistically expect a chance to get supercomputer experience before they graduate.

Industry is glad to hear it. For the first time, business analysts are beginning to note serious moves by corporate America into supercomputing. Though still wary of the considerable risk buying such high-priced hardware naturally entails, an increasing number of American companies are building in-house supercomputer power. The growing list already includes Boeing, Lockheed, GM, Ford, Exxon, Arco, Shell, Sohio and at least one film studio in Hollywood, which is starting to cash in on supercomputers' near-spooky talents in animation.

Industry's interest isn't piqued because supercomputers are getting cheaper; to the contrary, some businesses are willing to pay top-dollar for what they are just now beginning to see as vital investments in their futures, says Jim Andreason, marketing director for ETA Systems, Inc., a supercomputer manufacturer in St. Paul. Andreason sees commercial application as "the biggest growth area" in supercomputers through the end of the decade.

But for the time being, two major obstacles are preventing an all-out embrace by industry of large-scale computing, Andreason said. One is limited application. At present, the incredible computing muscle of such machines as the Cyber 205 is hobbled by a lack of instructions -- software --that can turn such enormous power loose on a wide variety of tasks. A supercomputer may be great at making models in theoretical physics, but it may just sit and blink when asked to design an artificial heart.

Secondly -- and fundamentally, Andreason says -- industry is already having trouble finding qualified people to run the few supercomputers they do have. American universities simply don't produce a large workforce of supercomputer-trained people. The bulk of such training is done by the federal government for new employees working in national labs, home to nearly half the supercomputers in the world.

"This is why programs like Florida State's are so very important to everybody," Andreason said. "Not only will you be developing new and creative ways to use supercomputers, you'll be producing the kind of people that industry has got to have to stay competitive."

Competition is driving both corporations and academe to scour the earth for the best computer-trained minds they can find. That's one reason SCRI, despite its youth, has already taken on a decidedly cosmopolitan air. O'Brien and Krishnamurti already have teams that could pass for delegations to the U.N. Just recently, two young Ph.D.s, one from Australia, the other from Israel, joined their groups.

"It's not that we're overlooking Americans," O'Brien said. "The problem is that in this country there just haven't been very many supercomputers for young people to get the kind of training we need. But, finally, that's changing." A "supercomputing infrastructure" is developing in America as government and industry step up their joint efforts to put supercomputers on campuses, he said.

Thanks to heavy governmental subsidies of academic supercomputing in Europe, and most notably in Japan, foreign countries have a jump on building such infrastructures. Few, if any, major Japanese universities are without ready access to state-of-the-art supercomputing. While some university-run supercomputer centers in this country are forced to charge users $2,000 an hour (and up) for time on their machines, the going hourly rate in Japan is as low as $60.

Duke, like O'Brien, sees a change for the better. "We're just now beginning to leapfrog ahead of the Europeans, for example," he said. "In the past, our researchers and students had to go over there. Now they're starting to come here. Right here to Florida State."

DR. T.N. KRISHNAMURTI, foreground, and students

Lannutti put down the phone and apologized for the interruption. "That was Roger Dashen, a physicist with Princeton's Institute for Advanced Study. He's coming to talk to us about how their new center and SCRI can work together."

The "new center" Lannutti was referring to is the John von Neumann Center for Scientific Computing, a $125 million supercomputing consortium soon to supply 12 universities with direct access to a Cyber 205. The center is named for the late mathematician and Princeton faculty member who designed and built some of the world's first computers.

Designed along lines similar to SCRI's, the von Neumann Center is one of four supercomputer research centers created last fall by the National Science Foundation. The federal agency plans to spend $200 million over the next five years in start-up costs for the Princeton consortium, for another 18-member consortium based at the University of California at San Diego, and centers at Cornell University and the University of Illinois at Urbana-Champaign.

When completed, the NSF centers will nearly triple the number of American universities with direct supercomputer access. Florida State, following the universities of Minnesota, Colorado State, Purdue and Georgia, was the fifth university to acquire a supercomputer. The University of Alabama is reported to be on the verge of buying one. Others seem to be leaning in a similar direction.

Competition is driving a search for the best computer-trained minds in the world.

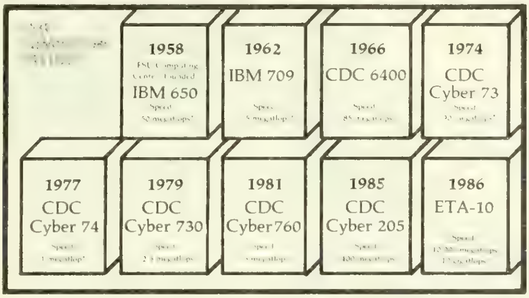
What does this proliferation of academic supercomputing power mean for SCRI?

"In the overall view it means that we are finally making substantial headway in a nationwide effort to help scientists and students become more competitive (through the use of supercomputers)," Lannutti said.

"At the same time, though, it means we've got our work cut out for us. One of our main objectives is to come up with new and better ways to use supercomputers. So that means competition in mind-power, in imagination. It's a friendly competition, but it's real, nonetheless."

When final hook-up is made, expected soon, SCRI will be part of a nationwide, satellite-based computing network, run by DOE's Office of Energy Research. The National Magnetic Fusion Energy Computer Network (MFE) as it's called, links ER users from coast to coast, via satellite stations at five regional locations, or "nodes." The closest network node to FSU is at Oak Ridge National Laboratory in Oak Ridge, Tennessee. A dedicated phone line will link SCRI's 205 to the MFE network.

Meanwhile, SCRI software specialists are already at work helping both campus and off-campus scientists write, or re-write, programs tailored to run as efficiently as possible



FLORIDA STATE'S COMPUTING HISTORY traces the evolution of computer speed available to researchers. For nearly three decades the campus has kept pace with innovations in computing technology.

Note: Speeds shown are theoretical limits. Graph indicates only the most powerful computer on campus in a given year, not total campus computing power.
** approximate*

on the 205. There is much talk of "vectorizing and optimizing" heard in the Keen Building. These are forms of mathematical massaging of data, necessary to take as full advantage as possible of the 205's hardware design ("architecture").

Matching architecture with cleverly written software is critical to supercomputers' speed and efficiency. No one needs a Ferarri, in other words, if the road's full of hairpin curves. Supercomputer programmers may speak in megaflops, pipelines, and parallelism, but they *think* in algorithms. These are mathematical "road-straighteners" -- tersely worded numerical languages that tell a computer how to get from here to there as fast as possible without wrecking the car.

Many experts say that if American supercomputing has any edge at all on the Japanese variety, its in the development of super-sleek algorithms. That's slim consolation, however, when one realizes that a Cyber 205 like FSU's is electronically capable of 400 million calculations per second (400 megaflops), yet commonly functions at less than a third that speed. The main reason? The algorithms have yet to be invented that can put all that horsepower to use solving real problems. The machines may fly through highly controlled tests at the factory, but when put to on-the-job work, like drawing a 3-D picture of an imaginary space vehicle, they're obliged to take their time.

Lannutti says that SCRI is dedicated to helping achieve a national goal of "200x", or computer speeds 200 times faster than what is currently attainable, by 1990. "It's going to come about through a combination of faster hardware and faster software. Our mission is to concentrate on the software."

he Age of Supercomputing may be upon us, but some scientists are greeting its arrival with skepticism. They are wary about what supercomputing may mean for them both personally and professionally.

The reasons vary, but a common theme is that supercomputers just aren't necessary to do the kinds of nuts-and-bolts research many scientists and engineers do. The argument is that supercomputers may be just dandy for modeling the ocean or designing airplanes, but they're downright clumsy at handling certain tasks that a minicomputer, or even a desktop, can hack through fairly easily.

Besides that, supercomputers aren't easy to use, skeptics say. In the jargon, they aren't "user-friendly." To use them with any skill at all, scientists must know at least something about how they operate, and that can mean (though not always) learning supercomputer programming. Many researchers are understandably reluctant to compound their workloads with such extracurricular activity.

"Scientists face the problem of not only having to know their disciplines well, but of having to know how to use computers, too," says Dana Hoffman, an ETA computer specialist. "Some, frankly, are overwhelmed. They say they've got enough problems without having to learn supercomputing. It's not hard to see their point."

One option now open to FSU scientists, Hoffman says, is to seek help from SCRI specialists such as Iris Young, an ETA consultant attached to the institute. "She can't write whole programs for people, but she can show them how they, or their associates, can write them to better suit their needs."

Still, there's a lot more to making a supercomputer friendly than just learning to speak its language. Ironically, now that access to the machines is better than ever before, many otherwise computer-literate scientists are passing up their chances to work on the machines, opting to work on far slower, but much easier-to-use mini- and microcomputers. They sit in their offices, enter their data directly into a familiar machine, and, in due time, get some answers. Such dispatch isn't the exception in supercomputing, but it's certainly not the rule. At least for now, Lannutti says.

"When people think of friendly computers, they sometimes name Apple's MacIntosh as an example," Lannutti said. "A person doesn't have to know much about computers to learn how to use a machine like that, which is why it's so popular. It's not unreasonable at all to think of supercomputers someday being used in a similar fashion."

A SCRI step in the direction of user-friendliness, he said, is a plan to design the 205's work-stations to be as convenient as possible. Technicians plan to borrow ideas from other supercomputer installations around the country and improve on them "as the budget allows," Lannutti said. The work-stations will put the power of the 205 directly into the hands of campus researchers. "A scientist will sit down at a terminal, ship his job out

MORE THAN 80 scientists from around the world attended a SCRI-sponsored symposium on high energy physics theory held on campus in April.

to the 205 at Innovation Park, have it run, and retrieve it onto a disk right at his desk."

No matter how cozy relationships between man and these supermachines may become, however, Lannutti, Hoffman and Jerry Stephens know that some scientists may forever shun the beasts. Some may never need that magnitude of computing in their research, while others will be content with conventional computers, Stephens said.

But supercomputers are such radically different academic tools that conventional ideas about how scientific research should be done are badly in need of "rethinking," he said. Both he and Lannutti feel that the machines pose a revolution in scientific thinking.

"In some areas, supercomputing has a long way to go," Stephens said

The machines pose a revolution in scientific thinking.

"But there's a lot of progress being made (in bringing supercomputing power to bear) on just about any science you can name. As these techniques develop, it will be the SCRI's job to show researchers on this campus, and elsewhere, how best to use the tool that we're fortunate enough to have."

"I hope you get the point of all this."

Lannutti was leaning forward, hands clasped tightly over papers stacked atop his desk. He had been talking -- to callers, aides and an interviewer -- for nearly two hours. Four gentlemen with briefcases had been waiting in an outside office for a half-hour. Patiently, one assumed.

"We're busy trying to make the best of a fantastic opportunity. It's right here, in our hands. We want to use it the best way we know how to contribute to the scientific and technological strength of this country.

"But we have another goal in mind, too. And that's to help Florida State. This tool gives FSU an opportunity to join the ranks of the top research universities in the nation. And that's truly an exciting notion."

He leaned back and smiled. Predictions?

"I expect that before too long there will be people giving talks at conferences around the world, saying they did their work with the help of SCRI. They'll say it's a good thing to work with us, SCRI, but hopefully, *us*, meaning Florida State.

"People will begin to see that we're a university to reckon with."

— *FRANK STEPHENSON*

PROFILE: THE CYBER 205

The Cyber 205, introduced in 1981 by Control Data Corporation, is likely to be the last supercomputer to bear the Control Data name. Since January 1, the marketing, distribution and servicing of the 205 has been the responsibility of ETA Systems, Inc., a company Control Data created expressly to carry on a tradition in supercomputer development pioneered by Control Data 21 years ago. ETA is designing the next generation of computers, the ETA-10, to be compatible with 205 software.

Thirty-three 205 systems are installed worldwide. Florida State's was No. 32. Several are in Europe, in universities and government labs in Germany and England. In this country, 205s are at work in industry (e.g. Chrysler Corporation, Sohio), in the federal government (e.g. Department of Energy, NASA, National Weather Service, National Bureau of Standards); and at the universities of Purdue, Colorado State, Georgia, Florida State and Princeton.

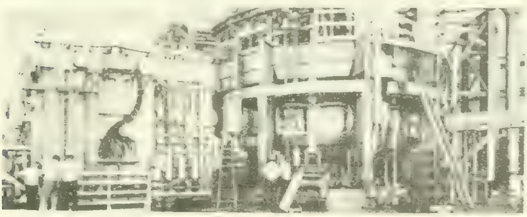
The starting point for every daily weather forecast in the U.S. is data compiled and analyzed on a Cyber 205 run by the National Weather Service in Suitland, Maryland. That machine digests more than 50,000 observations twice a day collected from ships, airplanes and satellites around the globe. When installed two years ago, that machine reduced the time required to make a two-day forecast from 22 minutes to 72 seconds.

With four million words in its central memory, FSU's machine is among the most powerful 205 systems in the world. Because it generates about 420,000 Btus of heat per hour, it is kept from melting by a 45-degree F. freon cooling system supplied by 35 tons of air conditioning equipment. This is enough cooling to make about ten 2,000-sq. ft. homes comfortable.

Next year, the machine will be replaced by the ETA-10, which is expected to be up to 30 times more powerful. That machine is being designed on a Cyber 205 by ETA engineer Neil R. Lincoln, who also served as Control Data's chief architect for the 205.

THE OFFICE OF ENERGY RESEARCH

Underwriting the Future



The Tokamak Fusion Test Reactor at the Princeton Plasma Physics Laboratory.

If DOE's Office of Energy Research ever decides to hang up its business, American scientists may as well forget about trying to:

- build a safe, reliable -- and near pollution-free -- fusion energy plant
- build a low-energy vehicle for mass transit
- stay competitive in the design and fabrication of new metals, plastics and other materials
- develop new and better energy sources
- find a clean way to burn coal, oil and gasoline
- safely, permanently dispose of nuclear waste
- find the nature of, and long-term answers for, environmental pollution
- figure out what matter is made of.

Though few may realize it, the nearly $2 billion spent annually by the office pays for research projects that effect every American's health, work, economic security, education and future. Without ER money supporting highly speculative, leading-edge research in thousands of federal and university labs around the country, American science and technology would be something other than No. 1 in the world.

If they didn't have to meet payrolls, pay taxes and fork over dividends, American businesses might be able to do more fiscally risky research. But there's not much consumer demand for a proton accelerator, an atom-smasher that can easily cost a billion dollars. Or giant, experimental devices to produce electricity from wind and waves. Or a nuclear plasma-maker, capable of heating gases to 180 million degrees Fahrenheit.

Fueling ER's fundamental research in fusion and nuclear energy, elementary particle physics, fluid dynamics, chemistry, engineering, computer science, mathematics, materials science, geology and environmental science is state-of-the-art computing.

That's a big reason why that when it comes to supercomputing, DOE is in a class by itself. With 21 machines now plugged into its various programs, including the Cyber 205 at Florida State, the agency is easily the biggest supercomputer user on the planet.

"There's no place in the world that's in our league as far as super computer use goes," says DOE's John Cavallini, an analyst in ER's Office of Scientific Computing. "We've probably got at least 20 percent of the (world) market."

DOE's supercomputers are clustered in eight centers around the country, including national labs at Oak Ridge, Tennessee; Las Alamos, New Mexico and in Berkeley, California. While some of the supercomputers in these centers are dedicated to classified nuclear weapons research, others are devoted to ER programs, none of which are classified.

Three of the biggest ER programs are in high-energy (elementary particle) physics, nuclear physics and fusion energy. About 90 percent of the country's research into the make-up of matter (high-energy physics) is ER-funded. This enormously complex, supercomputer-intensive research is primarily conducted in three government-run labs called accelerator centers. These are the Brookhaven National Laboratory in Brookhaven, New Jersey; Fermi National Accelerator Laboratory in Batavia, Illinois, and the Stanford Linear Accelerator Center in Stanford, California.

Multi-faceted investigations of atomic nuclei are carried out in seven nuclear accelerator facilities in government labs and on university campuses. About $160 million each year is spent by the ER on theoretical and applied work in low and medium-energy nuclear physics, heavy-ion nuclear physics, and nuclear theory.

Scientists are still years away from developing the technology to harness energy produced by atomic fusion. Unlike nuclear fission, in which atoms are split to release energy, fusion technology is based on what happens when two atomic nuclei combine. The sun and the stars, along with the hydrogen bomb, are examples of uncontrolled fusion energy. The process produces much more energy than nuclear fission and yet produces almost no radioactive byproducts.

Standing squarely in the way of scientists' efforts to control fusion is one of most profound problems known to physics: how does one build a container strong enough to bottle the heat of the sun? This year the ER will spend about $500 million in a continuing quest to find out. A good portion of that money will be spent for better computing power.

Cavallini said the ER has long been aware of the need for cooperation between the federal government, business and universities in expanding academic access to supercomputing. He said the agency is counting on programs such as SCRI to boost research in supercomputer computational science, seen by government analysts as the key to maintaining U.S. dominance in scientific and technological development.

Last year, ER's scientific computing staff established a program to monitor the supercomputing needs of university researchers throughout the U.S. An assessment showed that at current rates, by 1987 campus supercomputing demand will exceed supply by nearly four times.

Building the ETA-10

PROJECT SPEED

When it comes to computing power, these people are building a machine a breed apart.

They're building the soul — and body -- of a new machine

And what a machine it is bigger and faster than anything on the computing horizon. A cyphering *Wunderkind*. A pistol-whippin,' fire-breathin' *scorcher*. The closest thing yet to lightning in a can.

It's called the ETA-10. And if it's causing reporters to wallow in hyperbole, have mercy. By all accounts, the thing could be an abacus for the gods.

From the people who gave the world the Cyber 205, now comes a computer up to 30 times more powerful. If it's development continues at its present rate, its makers say the ETA-10 will be the fastest, most reliable computer on earth by the time it rolls through the factory doors, scheduled for late summer or early fall next year.

And grab your hat: the first one's coming to Tallahassee.

Florida State will host the world debut of the ETA-10, getting as much as a year's jump on other sites also slated to get the machine. In April, FSU was notified by ETA Systems, Inc., of St. Paul, Minnesota, that it was tapped to be the first on-the-job test site for what the company's 300-plus team of crack supercomputer-builders are hard at work on.

Formerly known as the "Cyber 250" or the "GF-10," the ETA-10 is, as far as computer-industry analysts can tell, the most advanced computer likely to appear anywhere in the world during the next two years. Even the most talked-about Japanese competitor is a soon-to-be-released machine rated nearly nine times slower than the ETA-10.

While FSU's acquisition of a Cyber 205 supercomputer has generated cross-country interest among the scientific community, the real excitement is over what lies just ahead. Built into SCRI's five-year cooperative agreement with the DOE and with industry is a powerful anti-obsolescence factor: as soon as new technology comes available, it comes to Florida State.

"That's the only possible way a research center of this kind can succeed," said Dr. Alvin Trivelpiece, director of the DOE's Office of Energy Research. "It's designed to stay current with the technology. Otherwise, you're missing the whole point of exploring the potential of supercomputing and in training young people with the best equipment available."

Power is the big pay-off in supercomputing.

What FSU's new test-site status (referred to in the industry as "beta" testing) means, then, is that the ETA-10 will join SCRI nearly two years ahead of schedule. That's great news for FSU scientists who are anxious to be the world's first users of the new machine. "We're keeping the ETA-10 in mind in everything we do," says FSU physicist Dr. Dennis Duke. "When it's here we'll be ready to take full advantage of it."

Beta-testing (alpha-testing is done at the factory) is the most crucial stage in the shake-down cruise of a new computer. Before such elaborately constructed machines can be successfully marketed, they must prove themselves as reliable, on-the-job workers, says ETA Systems' marketing director Jim Andreason.

"Beta-testing is a complementary arrangement. The vendor gets real-world testing of both its hardware and software. We get to find out from the people who'll be using the machine what problems exist, where we need to improve some things," he said. "At the same time, the user gets a jump on the technology. He gets the chance to put his hands on something no one else has."

ETA Systems, Inc. (the letters don't stand for anything) is a 1983 creation of Control Data Corporation (CDC) of Minneapolis. The brainchild of William Norris, CDC founder and chairman, ETA is built on a technological and financial base in supercomputing developed by CDC dur-

ing the past 21 years. In creating the CDC offspring, Norris said the company's "mission" was "to design, manufacture and market a 10 gigaflop supercomputer by 1986."

Translated, 10 gigaflops (hence ETA-10) is 10 billion calculations per second. In comparison, FSU's Cyber 205, among the most powerful 205 systems in the country, has a theoretical top-end of 400 megaflops (400 million operations, or calculations, a second). On top of that, the 205's internal memory (data storage capacity) is, at four million words, feeble compared to the ETA-10's 32 million. And that's just central memory. As designed, the ETA-10 will be able to call on a "global memory" of 256 million words. (Initially, FSU's machine will be configured for only 128 million words.)

As high-school science students know (or should know), power is defined as the rate of doing work. Making electricity dash through a computer's micro-innards in record time is one thing; saving time doing an ever-increasing amount of useful work in the process is something else entirely. Speed in tackling a big batch of work -- crammed into large memory banks -- equals power.

We've got a horserace in Minnesota.

Power is the big pay-off in supercomputing. A bulldozer operator may charge $75 an hour, but it can be downright scary to see what one can do in that time. Computers have grown steadily more powerful because the jobs in research and in industry have grown steadily tougher. At every step of improvement, computers have paid off in time saved, problems solved.

Ironically, though, they've also put human ignorance in ever-finer relief. The more we know, the more we find we don't know. Scientists and engineers take consolation in the fact that now they are at least *aware* of some things that they never knew existed. From studying natural behavior to designing a diesel engine, almost nothing is black-and-white anymore. Supercomputers are making the lines between the possible and the impossible blurrier by the day.

But following the bulldozer analogy, today's best supercomputers -- including the ETA-10 -- will be tomorrow's garden tools. Nobody knows that better than the people at ETA.

WELCOME TO ETA

THE NEW FORCE
IN
SUPERCOMPUTING

The glossy red-white-and-blue poster wasn't alone. Throughout the building, something was staring back at you from walls in offices, corridors, bathrooms. And with a message.

WE WILL WIN
THE
SUPERCOMPUTING
RACE

The message was becoming clear: We've got one heckuva horserace in Minnesota. And there's a bunch of folks in St. Paul betting they'll beat the world.

There's smart money aplenty saying they will.

Whoa! What's the rush? Is the sky falling? For Pete's sake, won't somebody explain why all these people are so lathered up?

"We've got a lot riding on the ETA-10."

Understated perhaps, but the reply more than matched a naive question. Jim Andreason of ETA went on to explain, in some detail, why otherwise well-adjusted young, and not-so-young people commonly spend their days "off" in the building, sitting in front of over-sized computer screens and staring for hours on end at the most distressingly complex electronic crazy-quilts one can imagine. And having a ball doing it, too. The situation had gotten clearly out of hand.

"We certainly don't have a morale problem, that's for sure. What's nice about these young people is that no one has ever told them they couldn't do something. So they do it. To them, its fun. It's hard work, but it's fun."

A stroll through ETA headquarters in St. Paul is like taking a ride on a time-machine. It's a backward glimpse into the late sixties, when super-bright youngsters worked in excited clusters around the country, feverishly trying to beat a clock started by a president who had told the world an American would walk on the moon by the end of the decade. By the summer of '69 they had made John Kennedy a prophet.

The analogy isn't all that farfetched. If not the purpose, the passion is much the same. ETA is in the race of its young life, and losing it could have serious consequences not just for the company, but for the health of U.S. science and technology. What's bad for the country's science and technology is foul medicine indeed for its economy and defense. More than a hand-in-glove relationship, it's closer to sinew and nerve.

Americans know all too well what being dependent on a foreign supplier of energy means to them personally. That's why the federal government has spent billions of dollars on energy research during the last decade, and plans to spend billions more. Today's oil-glut largely testifies to Americans' singleminded determination to conserve energy

Foreign supplies of oil are still necessary, though, for Americans to continue running their lives and businesses in the manner to which they've grown accustomed. For all it's value, however, oil is far less a prized commodity than is advanced ("high") technology. That's the stuff that put Neil Armstrong onto the moon, movies onto video discs, computers into wristwatches, cordless phones into cars, microwaves into millions of kitchens, gene-splicers into research labs, artificial hearts into a growing number of patients, and personal computers into more than 12 million homes.

Supercomputers get much of the credit. Without them laying a lot of its theoretical groundwork, today's technology would be a far cry from what it is. We would have yet to hear of a quark for example. NASA's space shuttle would still be on the drawing board. National weather forecasts would be good for only two days instead of four. Oil companies would have wasted a lot more money looking in the wrong places for oil and gas. And medical science would be oblivious to a technology capable of testing drugs inside a simulated human body.

But the goose that laid such golden developments at the doorstep of American firms and universities is in danger of starving to death. High technology feeds almost exclusively on the contributions of bright, creative minds. Science and engineering in this country is feeling the sting of foreign mind-power like never before. As if

America's foreign trade deficits weren't bad enough, overseas technological development -- especially from Japan -- threatens to make them even worse.

Federal officials, board chairmen and university researchers are very much aware of this disturbing development It's been nagging them for years. But only recently have they made any progress in doing something about it -- together. Last year, Congress funded Florida State's proposal to build SCRI, and soon thereafter voted $200 million more to step-up the pace at which similar programs are to appear on campuses during the next five years.

It's the stuff that put a man on the moon and 12 million computers into homes.

This is where ETA enters the picture, and the horserace. The company is one of only three U.S. supercomputer manufacturers, the others being Cray Research of Minneapolis and Denelcor of Denver. A whole new domestic market for supercomputers is opening up, which is triggering commercial and academic interest both here and abroad. Cray has been the world's largest supplier of supercomputers for several years. Denelcor, a much smaller company, is nonetheless a respectable contender. Both companies are hard at work building bigger, faster computers. Meanwhile, the Japanese are knocking at the door.

And ETA has yet to build a machine.

The company is hardly starting from scratch, though. If there's an ideal way to start a supercomputer company, ETA may have found it. Its close relationship to CDC makes it a formidable innovator in computer design. More than 130 former CDC computer specialists and engineers, people who built the Cyber series, are now ETA employees. They include Neil Lincoln, the chief architect for the Cyber 205; now for the ETA-10.

"We call ourselves the new force in supercomputers," Andreason said. "Give us a chance to get our machine out. Then we'll show you exactly what we mean."

Gentlemen, place your bets.

FRANK STEPHENSON

THE ETA-10

The machine being built by ETA Systems belongs to the most powerful class of supercomputers yet designed called vector multiprocessors. The phrase describes both the machine's software (the machine's operating instructions) and its hardware (architecture). By contrast, a Cyber 205 is a vector processor. A vector *multi*processor, such as the ETA-10, is essentially a multiple assembly of 205s.

Standard 205s are built with what is called LSI (Large-Scale Integration) technology. This means that its microchips contain large numbers of transistors. An average 205 chip contains around 10,000 transistors. About 175,000 LSI chips, arranged on 65 boards, make up the 205's "brain" or central processing unit (CPU). Thirty years ago a computer this powerful would have covered more than 125,000 sq. ft., or an area the size of two-and-a-half football fields.

The ETA-10 is a radically different machine. With up to 100,000 transistors each, it's microchips (which are just under a half-inch square) belong in the VLSI (Very Large-Scale Integration) category. About 300 chips fit on one ETA-10 board which is roughly 16 inches wide, 24 inches long and a quarter-inch thick. This single board, then, *is equal to the computing power of at least three, and up to five, Cyber 205s.*

The "multi-" aspect of the ETA-10 is achieved by arranging parallel configurations of two or more CPU boards. A full-blown ETA-10 machine will have eight boards. Each board will contain a mini-warehouse of its own memory -- four million words (the entire internal memory of a 205) stored in a cube about five inches square. Also, each board will access 32 million words of external, or "global" memory. A fully configured ETA-10, then, will command 256 million words of global memory.

ETA engineers are basing the ETA-10 on what is called CMOS chip technology. The acronym (for "complementary metal-oxide semiconductor") describes ETA's patented design. The chips, being made by the Honeywell Corporation, have a one-of-a-kind self-testing feature. A dot-sized, built-in microprocessor keeps constant check on the chip's circuitry. The feature allows each chip to be tested individually -- instantly -- either by an on-site operator or by someone hundreds of miles away. Once FSU's machine is installed, specialists in St. Paul will be able to troubleshoot each chip by telephone.

Unlike the 205, which depends on 35 tons of air conditioning to keep it from melting during operation, the ETA-10 needs very little heat dissipation. In fact, one ETA-10 CPU board (equal to at least three 205s) draws only the power of six 100-watt light bulbs. The machine will run perfectly well at room temperature... but only at half-speed.

Doubling the speed of the ETA-10 is accomplished by immersing the CPU boards in liquid nitrogen. This super-cold environment (-372 degrees F) make's the machine's circuitry superconducting, or able to conduct electricity twice as fast as normal. Once assembled, the entire ETA-10 will occupy about 70 sq. ft. of floor space, compared to more than 200 for the 205. It will consume in the neighborhood of 10 to 20 kilowatts per hour less electricity than the 205, and yet produce 30 times more computing power.

The ETA-10 is being designed to perform 10 gigaflops (10 billion calculations per second). FSU's 205 is electronically capable of 400 million calculations per second (400 megaflops). The fastest computer at NASA's disposal in 1969, when man first landed on the moon, was a Control Data 6600, a one-megaflop machine.

ETA hardware engineers say CMOS chips have major advantages over gallium arsenide technology as it now stands. Gallium arsenide is regarded by some as the chip medium of the future. Cray Research is basing its new machine, the Cray 3, on gallium arsenide chips. Control Data has conducted gallium arsenide research since 1979.

ETA says it is closely monitoring progress in gallium arsenide technology and will make a decision next year on how it plans to build its next generation of machines. The options are to go with gallium arsenide or a higher-density CMOS chip that can be changed out once improved gallium arsenide technology becomes available.

Work on the next generation has already begun. Tentatively named the ETA-30, engineers say it may be an "ETA-40 or 50" by the time it appears, sometime in 1992. They say that by then, a competitive supercomputer chip will contain around a billion-and-a-half transistors.

MEET THE SCRI

Launching Florida State's supercomputer program is a solid core of talent in science and technology.

ADMINISTRATION

DR. JOSEPH LANNUTTI,
Director

DR. JERRY STEPHENS,
Assistant Director for Administration

DR. DENNIS DUKE,
Assistant Director for Research

PATRICIA MEREDITH,
Assistant to the Director, Professional Services

SUSAN FELL,
Assistant to the Director, Administration

SCIENTIFIC RESEARCH STAFF

DR. ANNA HASENFRATZ
Specialty: lattice gauge theory
Education: Ph.D., physics, Eotuos University, Budapest
Native: Budapest, Hungary
Age: 28

DR. MIHALY HORANYI
Specialty: cosmic phenomena, astrophysics
Education: Ph.D., physics, Eotuos University, Budapest
Native: Budapest, Hungary
Age: 29

DR. KUO-CHENG LI
Specialty: supercomputer software development
Education: Ph.D., computer science, Purdue
Native: Taiwan, Republic of China
Age: 30

DR. GLENN SOWELL
Specialty: software development, computer algebra
Education: Ph.D., physics, Florida State
Native: Memphis, Tennessee
Age: 34

DR. CHRIS GEORGIOUPOULOUS
Specialty: supercomputer software for high-energy physics
Education: Ph.D., physics, Tufts
Native: Athens, Greece
Age: 26

TECHNICAL RESEARCH STAFF

SAM ADAMS
Specialty: support, evaluation of user work-stations
Education: M.A., chemistry, Auburn
Native: Columbus, Georgia
Age: 42

KEN HAYS
Specialty: hardware acquisition, evaluation, installation
Education: B.S., mathematics, Florida State
Native: Tallahassee, Florida
Age: 33

ERIC PEPKE
Specialty: hardware applications, communications
Education: B.A., computer science, Florida State
Native: New York City
Age: 23

JULIAN SOLLUHUB
Specialty: user training
Education: M.A., psychology, Johns Hopkins
Native: Tallahassee, Florida
Age: 40

JIM LYONS
Specialty: technical support for research
Education: B.S., electrical engineering, Georgia Tech
Native: Melbourne, Florida
Age: 39

DOUG LEE
Specialty: technical support, user needs analysis
Education: B.A., computer science, Florida State
Native: North Miami Beach, Florida
Age: 27

A TEAM INSIDE A TEAM

Most Americans know Hungary for its music -- the lively, haunting music of Liszt, Bartok and Kodaly. But Hungary is also famous for astute research in physics. Anna Hasenfratz and husband Mihaly Horanyi are part of that tradition.

The two have come to Florida State from Budapest's Central Research Institute for Physics to pursue the "professional opportunities" that SCRI offers. Though both are physicists, their fields are markedly different. Mihaly is exploring such cosmic phenomena as the dust tails of comets; Anna is interested in lattice gauge theory, which describes the interactions between elementary particles.

Anna says her older brother, also a physicist, helped spark her curiosity in the field. Her husband says he has always been intrigued by the sky. "Everyone is curious about what's up there. It doesn't depend on education. Space is really basic and of deep interest to everyone."

Though their three-year-old daughter is comfortable in her Tallahassee nursery school, and they are "very happy" to be working at SCRI, they admit to missing relatives, friends... and Hungary. "There's just something about living in an old European city," says Anna. "Yes," her husband said, "and the shops with large windows facing the street, filled with all the tempting goodies."

-- MARY TEBO

SUPERCOMPUTER USES

Supercomputer applications in industry, government and academe are expanding rapidly. Uses are being found for supercomputing power in such fields as:

BIOCHEMISTRY: Areas: protein dynamics; enzymology; large molecule behavior; genetic engineering; molecular biophysics

CHEMISTRY: Areas: catalysis; physical chemistry; polymer science; nuclear chemistry

ECONOMICS: Areas: economic analysis; forecasting; global economic modeling

ENGINEERING: Areas: robotics; aeronautics; fluid dynamics; nuclear technology; materials fabrication; computer design; automotive industry; quality control; instrumentation

ENTERTAINMENT: Area: animation

ENVIRONMENTAL SCIENCE: Areas: hazardous-material migration modeling; pollution monitoring and control

GEOLOGY: Areas: natural resource survey and recovery; earth science

INFORMATION SCIENCE: communications; signal processing; data management; control systems; retrieval systems

MATHEMATICS: Areas: basic research; computer science

MEDICINE: Areas: drug interactions; diagnostics

METEOROLOGY: Areas: weather modeling; forecasting; climatology

OCEANOGRAPHY: Areas: ocean floor mapping; current and wind modeling; El Nino forecasting

PHYSICS: Areas: elementary particle physics; nuclear physics; astrophysics; materials science; geophysics; molecular biophysics

Source: *U.S. Department of Energy, U.S. Department of Defense, Office of Technology Assessment*

Science moves, but slowly slowly,
creeping on from point to point ...
Yet I doubt not thro' the ages
one increasing purpose runs,
And the thoughts of men are widen'd
with the process of the suns.

— TENNYSON *LOCKSLEY HALL, 1842*

BIG COMPUTERS
A Place in Art, Letters?

"The giant brain" was what humanists called the computer in 1964, when IBM sponsored a conference to show that computers had a role to play in the humanities. Today, many humanists think of computers in the same terms, characterizing them as powerful machines that will enlarge the scope of human intellect and imagination. Thus it's no surprise that the supercomputer, a giant "giant brain," has stirred excitement at FSU in such disciplines as philosophy, music and art.

"I am pleased to have the supercomputer here, as are many of the people I've talked to," said Dr. Alan Mabe, chairman of the philosophy department. "I don't think there are too many people in the humanities who will have a use for it next week, but there are some things going on that may one day call for the power of a supercomputer."

Mabe was referring to such things as textual analysis, a method of identifying language patterns in texts; the computer-aided teaching of writing and logic; and the development of purely logical language, an endeavor closely tied to the philosophy of language. Though conventional computers have proven adequate for these tasks so far, Mabe says that these increasingly complex research in these areas may some day warrant supercomputing power.

In some ways, Steven Newcomb's interest in the supercomputer reflects Mabe's. Newcomb works with Dr. Jack Taylor in the School of Music's Center for Musical Research (CMR). He is interested in analyzing musical texts, or scores. To do this, he must have a computer to translate the musical notes, their time values, and every aspect of the written music into a more explicit computer language -- "absolute data," as he calls it. Then he can ask the computer to look for significant patterns in the music.

But for a computer to be able to pick up all the information on a written score, Newcomb says it must have "a huge amount of memory." Artificial intelligence -- a system that mimics human reasoning -- would be helpful in such work, he said. Hence, CMR is looking at the supercomputer with great interest. "If we can get an artificial intelligence language up," says Newcomb, "we'd like to see if it can read music and turn it into absolute data."

It can be just another tool for expanding expression.

Dr. Paul Rutkovsky is enthusiastic about the supercomputer not so much because of its analytical capabilities, but because of new possibilities it may present for artistic expression. Rutkovsky, an assistant professor in art, uses computer graphics to illustrate a satirical publication he produces called *Doodah Florida*. He says that institutes such as the Center for Electronic and Photographic Arts in Buffalo, New York have helped computer artists win academic respect.

"It's a natural process for an artist to use any medium that's appropriate to his images. I think of the computer as a tool to use both culturally and professionally." He regards the supercomputer as just another tool that could expand the boundaries of creative expression.

— *MARY TEBO*

SUPERCOMPUTING'S FUTURE TOPIC OF FSU CONFERENCE

Supercomputer users and policymakers from around the nation will converge on the Florida State campus June 9 for two days of discussions about where the U.S. is headed in supercomputer research and development.

Highlighting the conference will be a session of the U.S. House of Representatives Committee on Science and Technology, chaired by Congressman Don Fuqua (D-Fla.). The committee will open the conference June 10 by hearing three hours of public testimony from authorities in supercomputer design and application.

Sponsoring the event along with Florida State are Cray Research and ETA Systems, both Minnesota-based supercomputing firms; Peat, Marwick, Mitchell and Co.; the Westinghouse Corporation; and The Institute for Constructive Capitalism, a University of Texas-Austin research group founded by Dr. George M. Kozmetsky, executive vice president for The University of Texas System. Dr. Kozmetsky, a leading innovator of high-technology applications in industry, will conclude the conference in a summary address June 11.

The conference, to be held at the Florida State Conference Center, is open to the public. The agenda is as follows:

Supercomputers: A Key to U.S. Scientific, Technological and Industrial Pre-eminence

Sunday, June 9

2-6 p.m.	Registration at the Tallahassee Hilton; Tours of Florida State University
6 p.m.	Reception: The Governor's Mansion

Monday, June 10

8 a.m.	Buses depart Hilton for Conference Center
8:30 a.m.	Introductions: Dr. J.R. Kirkland, Florida Business Associates Welcoming Remarks: Dr. Bernard Sliger, FSU President Governor Bob Graham
9:30 a.m.	Hearing: U.S. House of Representatives Committee on Science and Technology Chairman: Congressman Don Fuqua
12:30 p.m.	Lunch Speaker: Sen. Albert Gore
2 p.m.	Conference Program Begins: Dr. J.R. Kirkland, Dr. Jesse Poore (Georgia Institute of Technology), Co-Chairs **Discussion Topic: Marshaling Resources for Supercomputer Technology: Industry Access to Basic Research and Human Resources**
2:15 p.m.	Federal Agencies: Research and Outreach Chairman: Dr. James Decker, U.S. Department of Energy Participants: Dr. Paul Schneck, Institute for Defense Analysis Dr. Charles Buffalano, Defense Advanced Research Projects Agency, U.S. Department of Defense Dr. Randolph Graves, National Aeronautics and Space Administration
3:30 p.m.	Break
3:45 p.m.	University Research and Training Chairman: Dr. Augustus Turnbull, Vice President for Academic Affairs, Florida State University Participants: Dr. Joseph Lannutti, Florida State University Dr. Glenn Ingram, National Bureau of Standards Dr. James Carmon, University of Georgia Dr. Ken Wilson, Cornell University Dr. Steve Orszag, Princeton University Dr. Larry Smarr, University of Illinois Dr. Sidney Karin, University of California-San Diego Dr. John Connolly, National Science Foundation
5:30 p.m.	Adjournment
6:00 p.m.	Reception Host: Tallahassee Chamber of Commerce
7:30 p.m.	Dinner Host: Peat, Marwick and Mitchell and Co.

Tuesday, June 11

8 a.m.	Buses depart Hilton for Conference Center
8:30 a.m.	**Discussion Topic: Strengthening American Basic Industry Through Supercomputer Technology** **Chairman: Dr. Donald Price, Vice President for Research, University of Florida** **Participants:** Chairman: Dr. Donald Price, Vice President for Research, University of Florida Participants: Dr. George G. Dodd, General Motors Peter Zaphyr, Westinghouse/Carnegie W.A. Ranck, Lockheed Bob Brauburger, Chrysler Edwin B. Neitzel, ARCO Dr. David Pensack, DuPont
10:30 a.m.	Break
10:45 a.m.	Commentary By Manufacturers Chairman: Dr. James Browne, University of Texas-Austin Participants: Tony Vacca, ETA Systems Brett Berlin, Cray Research Dr. Donald Fraser, IBM David Lowry, Denelcor Murray Scureman, Amdahl Danny Hillis, Thinking Machines, Inc.
Noon	Summary Address: Dr. George Kozmetsky, University of Texas System
12:30 p.m.	Buffet Lunch and Adjournment

PANEL 2 DISCUSSION

Mr. FUQUA. Thank you very much, Dr. Johnson.

Dr. Wilson, I guess the first time I ever heard the word as it applies to computers was from you several years ago. We have heard previous testimony from just about everybody that has made presentations this morning, starting with the first witness, Mr. Zanardelli, talking about hooking into smaller colleges—everybody cannot afford one. Dr. Good talked about the fact that the NSF could not keep funding more and more centers; that there would be a limit and we need to maximize the use of the centers.

Senator Gore mentioned about networking and how do we do that, making highways. Dr. Smarr mentioned about—in his testimony—about directing almost like when you have a reservoir with water, and you are directing it to other places—what is the best approach we need to take on trying to develop a comprehensive networking system?

Maybe all of you might want to participate in the answer.

Dr. WILSON. OK.

First, Mr. Chairman, I ask that my prepared testimony be part of the record.

Mr. FUQUA. We will make all of the testimony part of the record.

Dr. WILSON. I have been participating, in a rough sort of way, in the NSF networking effort as a member of the Networking Subcommittee of the Program Advisory Committee for the NSF, and of course I also see the same networking problems as head of the Cornell Center.

I mentioned that networking is extremely important for us. The most successful networking effort today is the effort of the ARPANET of the Department of Defense, and so I think one has to look at the way that program is organized to see how one should work on the networking effort on a national scale. And I think an extremely important part of the DOD effort was an adequately funded program with a few highly competent technical people running it.

The NSF has difficulties because it has not yet amassed a critical mass either of funding or people to make such a program go, plus I think it is extremely important that we not have a fragmented networking effort with one effort from the Department of Energy, another effort from NASA, another effort from the Department of Defense.

So I think this committee has got to work with all these agencies and, as I said, try to put in one place a group of people with the dedication to produce a national network for all those agencies, not just one, with the technical competence to deal with the incredible technical problems that the network issue faces. Any such effort is a research effort on networking; it is not simply an effort to put existing highly proven technology into place. You just cannot do networking that way.

Mr. FUQUA. Are you aware if the Office of Science and Technology Policy is involved in coordinating this, because it seems like a logical function for them to be very heavily involved in it?

Dr. WILSON. I have not seen a direct involvement by them. But I have only seen the effort in the NSF. I have seen from a distance

their negotiations with the Department of Defense about getting onto the ARPANET, but I have not seen a direct involvement by other agencies in terms of my own involvement.

Mr. FUQUA. Dr. Smarr.

Dr. SMARR. I want, as usual, to agree with Ken. I think it is an awesome problem; I think it is not just a governmental problem. I do not, in fact, know whether the Government has a big problem or a little problem as compared to business.

I was sort of stunned to find out General Motors, United Technologies have 10,000 installed personal computers already, and the personal computer was introduced—that they use—2 or 3 years ago. How are they going to hook up all their disparate researchers all over the globe together?

The telephone system is an earlier example of networking, electric power distribution system, water distribution system, highway system. This country has met this challenge many times before, and it will meet it again. And, in each case—railroads, any of these—there has been this difficulty in grappling with how the private sector and the Federal sector provide not only monetary resources to achieve this, but the regulatory resources, et cetera. And I think that this will be no less complex.

After all, the information—I think "network" is a bad word; I think the programs concerns an information distribution system—that an entire economy, everything that people do in their homes, is going to use, and probably—you know, TV and everything will be distributed on this network.

So it is hard to say what is a high enough volume to deal with this issue. I think in some sense there needs to be special hearings just on this topic, to bring private sectors as well as all these agencies together. I think your committee would be a perfect one to carry that out.

I think one of the important things the NSF is doing, though, is that in these national centers, each of the four centers are building different aspects of pieces of a network, and experimenting with what might work better, and building communities of users together to find what you like and what you do not like, and so forth. That seems to be real important right now.

The NSF is also funding experiments, say, where—they just funded the use of the BIONET technology to get up to a quarter million bits per second on some high planes. So, the universities, as always, are a great place to experiment in ways that companies might not have the resources to put to that—throwing the net that wide-open for possibilities. But then we have got to spin these off into the private sector.

I think, ultimately, to make the network—it is a very big problem.

Dr. KILLEEN. Well, I don't think there is always going to be more than one network. I do not think it is feasible to talk about combining these into one massive supercomputer network. Right now, there are several successful networks, ARPANET and MFENET, and I think the next step is to make these four supercomputer centers funded by NSF available as quickly as possible. And, of course, as previous speakers have pointed out, this is being done.

One thing that could be done, you know immediately, is to make dial-up access available over ARPANET, and that was indicated it is going to happen. One could also do as we have done and use a value-added network such as we use, TYMNET. NCAR, for example, makes their computing center available to scientists in the atmospheric sciences over TELNET, and there are several of these value-added networks that could be used for dial-up access to supercomputer centers. And that is a way of getting started right away, because anybody can call up a TYMNET number or TELNET. We have a TYMNET—it is called a machine, but anyway—in any case, it is a rather easy thing to do to get started.

Then one needs to build up these networks, as Dr. Smarr pointed out. And I think that it is a mistake to start talking about megabit rates and fiber optics and things of this kind right at the outset. The thing to do is to use that technology when you need it.

Mr. GORE. Will the chairman yield on that point?

Mr. FUQUA. Sure.

Mr. GORE. What is the—how many bits per second will ARPANET handle?

Dr. KILLEEN. I think their highest rate is 50 kilobits per second.

Mr. GORE. Fifty kilobits per second. But already there are a lot of projects which are pushing the scientists toward much higher transfer rates.

Dr. KILLEEN. We would like to have a higher transfer rate, of course. The thing is, you only use those in very peak periods.

The main reason to have a very high bandwidth is if you want to send graphic information, and we have learned how to do that by compressing the data. We have actually graphic terminals scattered around our network, and we are using 56 kilobit per second links, and the data is compressed and sent over, and then reassembled at a graphics user service station. And, obviously, these megabit things are very expensive. To connect all of these computers to all of these small universities with such links I think is a long, long way off.

Mr. GORE. Well, if the chairman will permit me to continue for a moment here.

There are several things which might lead you to a much higher transfer rate. Graphics is one, but multiuser projects is another. I mean, if you are really going to make good use of these supercomputers, you are going to have research teams operating simultaneously on a variety of endeavors. And if we are going to maximize the potential, then you are going to have each one of those research teams in virtually constant communication with comparable teams in other parts of the country; and you are going to use up that transfer rate over ARPANET, and the other alternatives you mentioned, very, very quickly, and it is going—to use the superhighway analogy again, I remember when you drive across country on those two-lane highways and, you know, it was just impossible. And now already the interstates in some places are being expanded to 12 lanes.

And it just seems to me—although, in the immediate future, for the next 3, 4, 5 years; I am sure you are quite correct—but by the end of the decade, and certainly the next decade, we are going to

be hopelessly constrained by these transfer rates in the current nets.

Dr. KILLEEN. I think we are in complete agreement. It is just that—my point is that I do not want to see these programs get bogged down waiting for these 12-lane highways when, in fact, one can in fact do things now. That is my only point.

Mr. FUQUA. Mr. Lujan.

Mr. LUJAN. Thank you, Mr. Chairman.

Just very quickly, although I do not know if there is a quick answer to it—it is an interesting panel, at least what I have gotten out of it.

Dr. Killeen, you are kind of interested in more capacity because those are your needs. And Dr. Smarr is talking about the two different functions at the University of Illinois. Dr. Wilson and Dr. Johnson are talking about getting it out and available to everybody else, access to these supercomputers by all these other universities.

I am wondering—well, first of all, the precise question: Do you look at providing that access even to the personal computer use?

Dr. SMARR. Yes. I think it is essential.

You know, the user interface you get on the few major personal computers is very nice. I mean, you—it becomes your little friend. You work with it all day long and you do some word processing and you do some spread sheets and some graphics, and so on. And that is why there has been such a revolution, 10,000 of them in GM, and so on.

Now, those cost $2,000. For $20 million, you do not get that. All right? So something is wrong, and I think it is easily rectifiable. And one of the things at Illinois that we are really dedicated to doing right away is to try to integrate the Cray, with the Cray time-sharing system the national lab has developed, into the personal computer so that, you know, you are sitting there and you have got the Lotus 1-2-3 spread sheet, and you have got the word processing here, and you have got the Cray here, and that is the way it should be. It should be there like any other tool. Those tools could be ported through these various windows.

I would like to be able to dump a big calculation into Lotus and use that to manipulate my numbers. Why do I have to write a Fortran program every time I want to draw one line? We have got to get our scientists' productivity up enormously if we are going to be able to do the job, and the personal computer has been a great tool.

I would like to see out of our development effort, you know, by the—say, next year some time—basically many homes in the country just hook their lap-top PC into the phone and start using the Cray. Obviously, you are not going to shoot these things across, but we do not really do that most of the time. Most of the time, we edit a few lines of code and run it again and check to see how the run went. That is why 56 kilobaud—you know, there are amazing things that are done at Princeton and the other centers using that network.

Mr. FUQUA. Just one quick question.

Should NSF try to direct each one of the supercomputer centers in a particular area? I know that is a difficult question because everything from the nuclear Navy to the War on Poverty Program,

everybody just says give us the money and get out of the way, that kind of——

Dr. WILSON. If I may address that.

One of the interesting things about the NSF centers is that they will be covering a very broad range of disciplines, and that makes those centers unique, not just among computing centers but among national centers of all kinds—contrasted, say, to a telescope facility, which is extremely important for the advancement of science, but it is used by astronomers only; or an accelerator facility, which is used by a physicist.

The supercomputer can be used by mathematicians, physicists, chemists, mechanical engineers, electrical engineers, chemical engineers, biologists, doctors, solar scientists, archeologists, architects, and the list goes on and on.

And when they complain about funding for our program, you have not to think of the $40 million, but you have to divide it among all those areas and then see how much funding you have got.

Mr. LUJAN. Does that mean it is so unique the National Science Foundation ought to step back and let you do your thing?

Dr. WILSON. Well, I just wanted to make that point, that that is part of the importance of the NSF centers, is that they will have that multidisciplinary character. And what is important about that is that when you try to divide up people according to how they should interact with the computer, you find that you will have one group of physicists that should be talking to mechanical engineers and electrical engineers; another group of physicists should be talking to chemists and biologists. And at Larry's center in Illinois and the Cornell center, especially, we will try to regroup our scientists and engineers so that the right people are talking to the right people, and that does not mean physicists talking to physicists.

But to come back to your question of specialization, yes, the centers will certainly develop specialization. Just for example, I have—my own expertise as director of the Cornell center is in the area of quantum physics, the area of how electrons behave, how subnuclear particles behave. And so you will see a certain specialization at the Cornell center in that particular area because I have the expertise to try to really push the program along in that area.

Larry's expertise will be in the area of astrophysics and very complicated fluid and gas flows. So, of course, you will see specialization in that area at Illinois. This is also important to the progress of science and engineering. But it will be specialization on top of a highly multidisciplinary character which we have not had previously in the supercomputer centers, and which is absolutely vital to preserve and protect.

Mr. LUJAN. Thank you.

Mr. FUQUA. Mr. Sensenbrenner.

Mr. SENSENBRENNER. Thank you.

I have a question for both Dr. Wilson and Dr. Smarr: How is the time at your center allocated—who is able to plug into the supercomputer?

Dr. WILSON. The present agreement is that as soon as we are up and operating as a national center, the NSF will allocate 60 percent of the allocable time to its programs. And that is going to be

very complicated. I do not know how they are going to handle it, because the physicists will argue with the chemists who will argue with the biologists to get that time. Forty percent of the time will be allocated to the centers themselves, and we have started an allocation committee and we will broaden it to national participation as we turn into a full national center.

Mr. SENSENBRENNER. Dr. Smarr.

Dr. SMARR. Well, all four centers are under this current agreement that Ken has just mentioned, the 60–40 split. I think you will see that evolve as things happen. But I think it is crucial that these facilities not be wasted by dividing them by the number of applicants so that everybody gets access. These facilities are built to solve problems that could not be solved heretofore, and that is very similar to why you build a Fermilab or you build a super colliding accelerator or why you build a very large telescope, because that machine can do things that otherwise could not be done.

Therefore, the allocation process cannot be democratic and has to be subject to a very strict peer review. But, certainly, at Illinois, with our 40 percent, we are going to try to keep large chunks of the time on the machine together to give to single researchers who have attained the high level of scientific peer review, who are doing the most ambitious, creative, frontier-type problems, so they can really bust the machine. That is the way you make progress in science.

When we went to Munich, you could not do anything like that, and it was only because we were intolerant as scientists—as scientists, to not being able to do our work—that we kicked the computer center where it hurt until we got the techniques to be able to do this sort of thing. As a director, I look forward to being kicked by the scientists who are going to say, "Your center does not make it; we cannot do this problem; and you had better get things better."

So, the allocation process will—there will be a lot of small users, OK, and they will be distributed widely——

Mr. SENSENBRENNER. Let me kick you once more.

Dr. SMARR. Good.

Mr. SENSENBRENNER. You used the analogy in supercomputers of a public water system where everybody who has a personal computer kind of has a faucet to turn on the water, and the people who run supercomputer centers are in effect pumping stations.

Now, I remember when I was growing up in Milwaukee after the end of World War II, which was a period of very rapid real estate development, that we had some real water pressure problems because there were lots of faucets and not enough pumps.

Is that what we are going to run into now?

Dr. SMARR. I think absolutely. I hope your committee keeps a weather eye on it.

Four years ago when this famous Press Report that three of us here were on—we pointed out the difference between capacity, which is what you are talking about, and capability, which is what the—the hardest problem that can be done on the system. So, an analogy could be what is the largest user, single user, in the community, for water—one company or one individual—versus the need to distribute to all members of the community.

Those are quite separate issues, and yet we have got to grapple with both of them. That is why I argue that the supercomputer centers cannot be thought of as the only generating capacity in that network. The network—one of the reasons the network is there is so that various nodes of the network can develop where there is marketplace demand for their special niche.

I think you will see nodes on the network who only want to shift cycles and do a very good job of it and keep everybody happy changing that cycle, and that is a very critical thing to be done. One of the good points about these national centers is that the marginal cost of adding a second supercomputer is small. The supercomputer is maybe a quarter of our budget at one of these national centers. The rest of it is all the professional staff, the buildings, the software, all of that sort of thing.

To add just one more Cray, for instance, to our facility is a fairly marginal cost, and it would double the capacity. This is one of the issues that has got to be dealt with. And if you try and get everybody wanting to use the supercomputer, you will not see breakthroughs in science that we hope to bring to you with the centers. So there has got to be a balanced program.

Mr. SENSENBRENNER. Thank you.

Mr. FUQUA. Mr. Boehlert.

Mr. BOEHLERT. We are where we wanted to be a couple of years ago. We have got the four centers announced and up and running. And, Dr. Wilson, you will be fully operational in the fall, you anticipate.

My question is this: That 2-year program we talked about of roughly $7 million a year for NSF—and you promised great things, and we have great expectations—should we accelerate that funding now? Should we be satisfied with the 5-year program? Should we try to compress it? Should we try to increase the dollars that are coming to you?

I want you to know he did not plant that question.

Dr. WILSON. To put it into perspective, the proposal from Cornell to the National Science Foundation requested $68 million over 5 years. The actual award was $21 million over 3 years. Now, obviously a certain amount of programs got lost.

The other things we are dealing with are very complex issues. For instance, if I were running an accelerator program such as we have in the storage unit at Cornell, I could come to you and say, "Here is our budget. Here is what we need. Please fund the difference." This is much more complex.

I am not sure whether we can make it on the $21 million, but I am not sure if we could even make it if we had the $68 million. I mean, the program at Cornell involves incredible risks on the funding side. You know, just for example, we are waiting for a computer—which I cannot talk about—which will be fantastic. But it might never appear. And no matter how much funding you get, it might never appear. It might also appear, and there be no dollars at all because the company is very excited, and we are talking about having it and not having to pay anything for it.

We talked about capacity, can we meet the capacity? The NSF has told us recently that the demand on their present resources is growing very rapidly. Under normal circumstances, you say well

we have to start controlling allocations, we have got to cut people off so that the best people can get the time that they need. But in this program, I do not believe that we have to do that. I do not have to come up to you if we cannot meet all our needs and ask for money, I go to IBM, I go to Floating Point Systems and try to set things up so I can go to companies in oil and automobiles and glass and chemicals and say "Look at the great things that are happening here, why do you not help?"

And I am going to the computing industry and saying, "Look, there is a fanstastic need for a supercomputer. Why do you not quit producing them by hand, one a month, and get an automated assembly line which cranks out a thousand next year, because that is what we need and make them so that they do not cost $5 million each?" And because supercomputing is such an incredible idea, the industry can do that if they want to, and I am trying to make them want to, just as hard as I can.

Mr. BOEHLERT. Dr. Smarr, you talked about the marginal cost of adding a second computer at one location, which would seem to argue in favor of rather than adding new university research centers like the four we have agreed to, that we should add to what we already have rather than to supplement with new centers. Is that what you are suggesting?

Dr. SMARR. I think that under the constraints of the budget voted by your committee, there is no choice but to augment, to fund the current four centers and then get on with building the network, and I do not see enough money in the budget to really do that. I think the original $100 million a year proposed was a good, conservative estimate of what this Government's commitment should be in this area. I think it is what this committee recommended last year in authorization and I think that was a correct assessment of the need. To the extent that we have less than half that, we have less than half the program and there is no way around that.

The industrial dollars which we are seeking very actively, and I think all the centers are seeking very actively, are not going to replace those dollars of basic research in universities. They are going to be involved in the program, but they are going to be largely for applied research which is very exciting, but it is not the same as basic research.

So I think those industrial dollars will leverage your investment enormously in generating a greater GNP out of which everyone will benefit, but it should not be mistaken for—I mean any scientists, any university, you are our only hope—that Federal agency which funds basic research at universities and that is it. And so you know, we have to protect the NSF and we have to protect our program because there is no alternative except for augmentation but not replacement.

Mr. BOEHLERT. Dr. Wilson, you mentioned in your testimony that the gap at Cornell—I think in your exact words, you are not signing up enough industrial users. I know you are trying, what is the problem?

Dr. WILSON. The difficulty is the following—the computing industry is very enthusiastic about working with the universities and you know we have incredible support from IBM, from Floating

Point Systems and other companies as well. I mentioned that we are talking to Gould, which is based here in Florida, but when you get to dollars and cents, when you are talking to the computing industry you are looking at their marketing budget.

They are coming into the universities because it helps them sell computers and other reasons as well. I mean, IBM is a very science and engineering oriented company and that is a very important part of their involvement, but when you come to industrial users, the connections that they establish with universities are at the very end of a very long pipe because you go from the marketing budget to their basic developing budget for next year's products to the advanced engineering development, to their advanced scientific research and by the time you get to the advanced engineering and scientific research, there is a very small funding pot even in a company like General Motors.

To ask them to take part of that pot to support a university, it is very hard for their own scientists and engineers to support that because that is their money and there is not very much of it. So establishing a climate where industrial users participate in our program is much more difficult than establishing a climate for the computing industry to participate. Yet it is incredibly important for their future and ours. I am devoting a lot of my effort to trying to get the message across that despite the change, despite that problem of money they have got, it is important that they all get together.

Mr. BOEHLERT. Do you find it easier than it was 2 or 3 years ago?

Dr. WILSON. Well, I think I have been building it up for a long time. I think it is progressively getting easier for me to make a pitch but you never know when you have succeeded until you succeed. And I hope that will be soon, but I am not really able yet to tell you, you know, when we are over the hump.

Mr. FUQUA. Your time has expired.

Mr. Boucher.

Mr. BOUCHER. Thank you, Mr. Chairman.

I only have one question but I would appreciate each of the four panel members responding. Realizing that the National Science Foundation supercomputing budget is somewhat less than half of the original recommendation of $100 million annually, what are other countries doing today in supercomputing, and are we in fact being competitive with a national effort that is less than what this committee recommended and what Bardon and Curtis recommended, and if we are not being competitive, what should we do to make sure that we are?

Dr. SMARR. Well, I think that it is just not appropriate to let other countries tell us what to do. I think we can set our own agenda. I think this committee has done that. I think it is an extremely good agenda, I think it has all the right components. You have got to get these supercomputers into a few centers so we can train the students and faculty and develop the software and learn how to use them so that everybody benefits from that and benefits from doing it. We do have enough money in the out-years for those programs to really succeed and stay at state of the art.

We have got to run this program up by some amount, not a factor of two, but some, you know, marginal amount has got to be

available or else we are not going to be able to do what we said we would be able to do.

You have got networking set out and you have already heard from Dr. Good that these plans 3 years from now for fiber optics going in, there is a lot of work going on there but not nearly enough.

You have got the idea of what I call the sanctity of the local node—you cannot say because these are supercomputers, you can take away the computers from people locally who built up around minicomputers. If anything, they are going to need more computing power locally to handle this burst of information from the supercomputer. And so I think that part of the program needs to be strengthened both in the programs of the NSF as well as in the way of work stations and small computers. That is the very purpose.

I think if we can simply get back to that $100 million, I think we can do that program and I think you will not be hearing any more about the foreign competition. I think the Japanese program is a good one, the German program is all right, it is not as ambitious as the Japanese one in some ways, but we should be setting our own agenda.

Mr. BOUCHER. Let me just ask a question. Are you concerned that unless we meet the $100 million per year recommendation, that we run the risk of surrendering our lead to the Japanese program?

Dr. SMARR. Absolutely, and for just an enormously tiny amount of money. I mean, $50 million in a $4 trillion a year economy? It is scandalous that that would stand in the way of the entire future of this country as an information society, it is a scandal. But we have to deal with these things and we would like to help you but you know I think if history is written so that $50 million is invested in the bottom of this research pyramid and our country did not make it, that is a very sorry record.

Mr. BOUCHER. Dr. Wilson, do you concur?

Dr. WILSON. I cannot put dollar amounts on it but I think what I would say is the following: I think it made an incredible difference to the program to have that $40 million instead of $20 million. I mean, I never encountered a situation where, you know, that small a change made such a difference. And I think if we had started with $20 million we would have almost had to tell you today that the program is a failure.

With the $40 million, we have got an excellent start, but the crucial thing is not to let this initiative die, not to squeeze it down. I sit on that Allocations Committee, there has been so little access to supercomputers up to now, that now is not the time to start talking about "No, you cannot have time on the system."

The people who are crying today for supercomputers are not the—maybe a few years out we can start telling people "No, your idea is not good enough, go back and rethink it," but today is not the time to do that and I do not want to see us turning anybody down until the present users of supercomputers have grown to the point that the people coming in, you know, the people applying today for supercomputer time are real pioneers. By telling people "No, you cannot go west," we don't want to do that. And so it is

going to be extremely important that the funding for the program increase at a level where we do not destroy the excitement that presently exists.

I think people have come over here from Germany because they see this is where it is happening right now. Let us not destroy that spirit.

Mr. BOUCHER. Dr. Killeen.

Dr. KILLEEN. Well, right now I would say that the United States is not behind Japan and Western Europe in any of the things that we have been talking about. As far as supercomputers in Western Europe, they actually come from this country, and of course as was point out by Larry Smarr, they made them available to a broader discipline earlier than we did. I mean, there are several centers and so on in Germany and there are several Cray's available to the universities throughout Great Britain, for example, and they were doing that before we were, but I do not think their networks are any more advanced.

As far as the Japanese are concerned, the Japanese universities get the very first supercomputers that are made from various companies. They go to the universities. They are not necessarily used by other universities however. It is not a network situation in quite the same way we are talking about. The fusion program in Japan has links to other sites and it also has links to our center as well, and they have computers there which are comparable to our XMP's, but they are certainly not ahead in my opinion. But I think that the earlier statements that the momentum should not be lost, is crucial here.

Mr. BOUCHER. Do you feel that an additional sum of dollars, something beyond the NSF's current supercomputing budget, will be necessary to keep that critical momentum from being lost?

Dr. KILLEEN. Yes, I do. I think that the network is going to require quite a bit of money and I do not think there is enough money now to fund the supercomputing centers themselves and to develop networking.

Mr. BOUCHER. Dr. Johnson.

Dr. JOHNSON. I think I echo what my fellow panel members have already said and one thing we must keep in mind is to keep the supercomputer centers within the universities because where better can we get the training that is needed and the expertise to develop the software as well as some of the hardware capabilities.

We keep talking about four centers. At Florida State University we do have a center funded primarily by DOE and I want to emphasize this once more—it may have been lost in the shuffle here—they are supporting the research institute, not the computing itself. So we do have, like the NSF centers, we have 35 percent of our time available to us. We are working with industry, we are working with other universities, particularly in the southwest.

Mr. BOUCHER. Thank you.

Mr. FUQUA. We thank our panel.

Senator Gore, we are going to take our time for lunch, so we do not have to interfere with our time.

We want to thank our panel for joining us this morning, we appreciate it very much.

While they are departing, I have been calculating. While we have been in this hearing this morning, almost as many people have moved to Florida as are in this room.

[Questions and answers for the record follow:]

University of Illinois
at Urbana-Champaign

National Center for Supercomputing Applications

153 Water Resources Building
605 East Springfield Avenue
Champaign
Illinois 61820

Dr. Larry L. Smarr
Director

217 244-0072

ANSWERS TO SUPPLEMENTAL QUESTIONS

TO CONGRESSIONAL TESTIMONY

OF JUNE 10, 1985

QUESTIONS SUBMITTED BY MEMBERS OF THE

SUBCOMMITTEES ON

ENERGY DEVELOPMENT AND APPLICATIONS IN SCIENCE

AND ON

SCIENCE, RESEARCH, AND TECHNOLOGY

OF THE

COMMITTEE ON SCIENCE AND TECHNOLOGY

U. S. HOUSE OF REPRESENTATIVES

DON FUQUA, CHAIRMAN

ANSWERS BY

LARRY L. SMARR

DIRECTOR, NATIONAL CENTER FOR SUPERCOMPUTING APPLICATIONS

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

OCTOBER 28, 1985

QUESTIONS FOR THE RECORD

1) Should NSF manage the new national centers as a single resource or a collection of independent capabilities? Why?

ANSWER:

The NSF has created three national centers to deliver commercially available supercomputing resources and one experimental center to develop new computer architectures. There is a great diversity in the equipment that is available at these centers, as well as in the strategy each center has for using it. Furthermore, in certain cases, the particular equipment and local expertise combine to create natural strengths which are ripe for in-depth exploitation at that center. Examples are, medicine in the San Diego center, astrophysics in the Illinois center, and advanced graphics at the Cornell center.

Therefore, I would advise that while we are still in the experimental stage of determining how best to provide advanced scientific computing resources to our nation's researchers, we should allow the centers to develop as independently as possible. I would oppose any strong control from Washington over the approaches the centers take, subject, of course, to annual review by the scientific community and the National Science Foundation.

The NSF has been very enlightened in this respect. It called for a national competition to create as many models as possible from the research community. It has encouraged many of the creative ideas that the community has suggested. The only demand that the NSF has made of all centers is that they allocate resources on a national basis, with excellence, not institutional affiliation or field of study, as the criterion for access. I think that after several years of independent operation of the four centers we will have learned a great deal about how to best allocate scarce computational resources to a diverse research community.

2) The NSF Bardon-Curtis Report projected the need for a NSF Advanced Scientific Computing program to be funded at $100 million per year. As the new NSF and DOE centers begin to train potential supercomputer users and new scientific problems requiring supercomputers arise, the demand for access will no doubt increase. What do you see as the long-range need for national supercomputer centers, mini-supercomputer centers, and local workstations?

ANSWER:

The far-sighted Bardon-Curtis Report of the National Science Foundation correctly saw the need for a $100 million per year program in advanced scientific computing. Unfortunately, because of various budgetary concerns in Congress, the program is currently funded at less than half that amount. This much-reduced funding level is going to make it difficult for the NSF and the universities to fulfill the vision laid out in that report.

2) ANSWER (cont.)

The inadequate funding is going to hurt in at least three areas. First, the centers have been given "bare bones" budgets with which to build themselves. The staffing levels, while adequate, are extremely thin and do not leave much room for experimention in new technologies. Such experimentation will be crucial if we are to understand how both universities and industry can best use computational resources. Second, there is not nearly enough money to develop a true national network, which would give researchers in all universities quality access to the supercomputing centers. Third, there is insufficient funding for computational power at the researcher's home base. Researchers are going to require more local computational power than they have now in order to turn the increased flow of raw data from the supercomputer centers into scientific knowledge.

As the question points out, we will experience in the near future a quantum jump in the number of students trained in the use of these supercomputers. Furthermore, faculty and staff at universities will modify their research goals so as to be able to take advantage of these new computational tools. This will mean a much enhanced growth rate of demand for computational resources. This enhanced demand will probably grow more rapidly than the supply of cycles planned in the current budgets of the four national centers.

I believe the first priority should be to build a national network to provide supercomputer access to a major fraction of the university research community. Once this is done, we can tackle the problem of how to distribute increased computational resources on that network, so as to meet the rising demand from all over the nation.

I recommend a three-pronged approach to adding additional local capacity on the network. First, we should place more capacity at the four national centers to meet the general increased demand. There is a considerable economy of scale in adding supercomputer capacity to an existing center once the buildings, staffs, and the support equipment are all in place. Second, as the requirements of a particular research program grow to where they become a substantial fraction of a national center's supercomputer, then these programs should acquire their own "mini-supercomputer". Thus, local <u>production</u> capacity should be selectively added to the national network. Third, local <u>interactive</u> computing capacity can be increased by adding individual workstations on the researchers desk. These versatile scientific workstations are becoming essential for analyzing the supercomputer output.

I do not believe that the current NSF budget can adequately fund the national network or this three-pronged approach to meeting the enhanced growth rate of demand on that network. Therefore, I strongly urge this committee to bring the Advanced Scientific Computing budget to the <u>$100 million per year</u> figure advised by the Bardon-Curtis report.

3) (a) How do you view the way in which NSF currently allocates supercomputer access time to researchers?
   (b) What impact will the current allocation of time at your Center have on its mission and the way in which its products and services are managed?

ANSWER:

(a) The current NSF policy on allocating supercomputer access to researchers requires that 60% of the cycles of each of the national centers be allocated in Washington by the NSF program officers, while 30% is allocated locally by a peer review board composed of nationally prominent scientists and engineers, and 10% of the time is allocated on a full-cost recovery basis to involve industries in joint research projects with university personnel.

This decision represents a compromise between the legitimate desires of the program officers to keep track of the overall research programs of scientists or engineers in their area and the need of the supercomputer center to make sure that its particular capibilities are used in an effective way. I think the compromise will let us get started, but I consider it to be an arbitrary and artificial division of a center's resources. I expect this division to evolve with time as we gain experience.

My belief is that science would be served best by putting all of the allocation in the hands of the peer review board at the Center. The members of the board are closer to the Center and better understand how to match the Center's strengths to the problems submitted by researchers. Such a policy would also be more in keeping with the allocation policies used by other national centers with specialized resources, such as telescopes and particle accelerators.

(b) The negative impacts of the current allocation policy on our Center are three-fold. First, if a Center peer review board is only going to allocate 30% of the cycles, it will be rather difficult to get the best quality researchers to volunteer their time to be on such a peer review board. Secondly, it sets up two different channels for getting onto the machine. This will surely lead to a redundant effort in the peer review process. Thirdly, it reduces our ability to focus on particular scientific problems which can only be solved by using the large amounts of computational resources which our Center is designed to provide.

I believe that both the NSF and the Centers will be flexible on these issues and that within a year a better allocation policy will have evolved.

4) (a) Who establishes the policy for the allocation of supercomputer time at your Center?
   (b) Are research proposals for large segments of time, versus proposals for normal cycles of time, evaluated by the same criteria?
   (c) Are the criteria used by NSF to allocate time at the NSF National Centers compatible with needs of all disciplines, especially interdisciplinary research proposals?

ANSWER:

(a) The Chancellor of the University of Illinois, Thomas Everhart, has invited some 30 major research universities, distributed nationally, and 10 major research corporations to send a high-level representatives to participate in an Advisory Council to the National Center for Supercomputing Applications. The first meeting of this group was held September 9 & 10, 1985.

This Advisory Council is the senior policy forum for how our center can best be utilized by the national scientific and engineering community. It has the great advantage of having scientists, university administrators, and corporate leaders involved in a single forum. Many of the individuals on this council will have overseen the entire research enterprise of their university or research laboratory in industry. The University of Illinois has asked this council to develop a slate of candidates of eminent researchers to serve on a peer review board that will allocate the current 30% of cycles.

(b) The request for large amounts of time, for example over 250 hours of supercomputer time per year, is the central problem in allocation. It is essential that a number of these large problems be run, since it is these problems that really require a supercomputer. However, since there are only about 8000 usuable hours in a year, there obviously cannot be too many large problems run per year. Thus, review will have to be much stricter for large jobs than small.

At Illinois we have also developed much expertise in making large jobs run more efficiently by making changes in the computer programs. We believe that anyone allocated such large amounts of the resource should show evidence of code efficiency before they consume the scarce resource. The experience thus gained will also be used to make smaller jobs more efficient as well.

(c) I believe that there is a problem with the current NSF allocation process in that the education of program officers about the uses of advanced scientific computing resources varies considerably from area to area within the foundation. This means it is rather difficult to get an consistent division of time between different subdisciplines.

Furthermore, from year to year the subfield of science or engineering in which breakthroughs can best be made by applying the tools of supercomputing will vary. Therefore I think it is inappropriate to give each NSF sub-division such as Physics, Chemistry, Civil Engineering, etc. a particular percentage of the total resource that they are "entitled to" each year. I think it is much better to have a more fluid circumstance where the focus rotates from field to field over a period of years as the science requires it to.

53-831 O -86 - 7

4)(c) ANSWER (cont.)

Finally, a number of areas, particularly interdisciplinary research proposals, may "fall between the cracks" of the current divisional structure at the NSF. One of the most exciting things about the use of modern supercomputers is the bringing together of investigators who previously would have considered themselves in totally disparate fields. For instance, the mathematics of Monte Carlo, which accounts for many of the large supercomputer requests, is common to fields as far apart as elementary particle physics, nuclear physics, solid state physics, chemical reactions, and biophysics. Here at Illinois in our Materials Research Laboratory we have already seen researchers from many different departments coming together and sharing their common computational approach to their varied problems. However, in other universities these same investigators would typically have to send grant requests to quite different subdivisions in the NSF.

5) You underscored the importance of supercomputers to industry as well as industry's contribution to the NSF National Centers. The major interest of the private sector will no doubt be technology development and application. Do you feel the time will come when the National Centers will be self-supporting, or will federal support always be necessary for experimental basic research?

ANSWER:

We at Illinois have been very gratified by the strong interest shown in our supercomputer center by a number of major corporations in industry. It is true, as you state, that the private sector is interested in some of the new technologies that we are developing and new applications which are being developed by basic researchers. However, the interest I believe goes much deeper than that.

What the corporations are more interested in is having their best researchers and computer managers working side-by-side with the best researchers in the universities. They feel that this sort of sharing will be mutually enriching in ways that are impossible to completely determine at this time. Surprisingly enough, most corporations have said that they have little interest in just "buying cycles" at the university centers. They feel that they can quite adequately handle this either with their own machines or by buying computational services from the private sector.

At Illinois we are attempting to expand our supercomputing center to include a new supercomputer, with an associated research center for industrial applications, which would be used with our corporate partners. This would be located next to and interact with the center for university researchers that we are funded for under the NSF grant. Thus, we see the industrial support as leveraging, rather than substituting for, the federal support.

We believe that this new kind of equal partnership between the Federal government, the state government, the universities, and the corporate world will be extremely beneficial to all those concerned. Each component is contributing a unique characteristic to the partnership based on its legitimate mission. It would be unwise for the Federal government to decrease its support for basic academic research and thus weaken that vital part of the partnership.

5) ANSWER (cont.)

Therefore I believe the national centers must rely on Federal support for the forseeable future. However, these centers have the opportunity to become the nucleus of enlarged centers in which the fruits of the investment made by Congress can become more directly shared with the American economy.

6) (a) What do you feel the federal role in support of supercomputer research, training and application should be?
   (b) What should NSF's role be?
   (c) In your view, are supercomputer policies and technology issues adequately coordinated among federal agencies?

ANSWER:

(a) I believe that the Federal role as advanced by this committee is quite appropriate. It views itself as funding the sort of basic fundamental research which industries may not judge to be within their mission. For example, the new computer architectures which are being supported by the NSF, DOE, and other Federal agencies are exploratory basic research in determining what types of multiprocessors, memory hierarchies, etc. will be useful. Once these ideas are developed, I believe these architectures will move correctly into the private sector, where companies will use the market place in its traditional role for allocation of resources.

Finally, the universities have always been the primary source of training and basic research, both for graduate students and faculty. The initiative by this committee strengthens that traditional role a great deal, which can only benefit greatly the other elements of the American society.

(b) The National Science Foundation's role should continue to be what it is now: the Federal agency which supports basic research in American universities. While there are other Federal sources for such funds (such as DOE and DOD), these often are focused on priorities that those agencies have rather than on the sort of broad, undirected research approach that the National Science Foundation has historically supported.

Therefore, I believe that the NSF's role should be to:

1) furnish adequate resources to the four new national supercomputer centers,

2) take the lead role in building the national network, hooking together all researchers in universities with the national supercomputer centers, and

3) supply an adequate amount of computational resources at the researcher's home base so that he may take advantage of the supercomputing facilities.

6) ANSWER (cont.)

(c) The coordination among Federal agencies is never as great as one would hope for. There has been, however, a good deal of direct interaction between some of the NSF supercomputer centers and the DOE weapons laboratories, resulting in considerable technology transfer. A notable example is the use of the Cray Time Sharing System, developed in DOE labs, by the Illinois and San Diego supercomputer centers. This sort of direct center to center interaction is probably more useful than attempts to establish such coordination between the agencies themselves.

7) Clearly, other universities will want their own supercomputer.
   (a) Is there a hierarchy of centers needed or a variety of capabilities, i.e., by discipline, region, performance, level, etc.?
   (b) What policy issues should be considered by federal agencies to provide for future national needs, coordination of federal and non-federal centers, and cost benefits such as economies of scale?

ANSWER:

(a) One should recognize that a number of universities besides the four NSF national centers currently do have their own supercomputers. Certainly nothing should be done to discourage universities from using their own resources to acquire such machines if they deem this to be important. However, I believe the difference between the four NSF centers and the individual ones at universities is that the NSF centers are meant to be permanent <u>national</u> centers, with a national constituency. There is very little time on the machines reserved for the home university's researchers.

A major policy question facing the NSF and Congress is whether supercomputer centers should be created for specific disciplines (for instance, in connection with certain astronomical observatories or particle accelerators). Again, I think that the best way may be to allow the system to evolve for several years and to see how the ingenuity of people to solve their problems with the NSF supercomputing centers develops. My feeling is that those disciplines which might require supercomputing capacity ought to use their scarce financial resources to develop their "value added product" (e.g. image processing, specialized data bases, etc.) to the network and get their cycles from the centers which have the economy of scale to produce the cheapest computational power.

The national network should provide sufficient bandwidth for supercomputers at non-federal centers to communicate effectively with supercomputers at other centers. Ultimately, I believe there will be a large national "menu of computational options" confronting a researcher. He will have developed his favorite supercomputer to run his problem on. It is the purpose of the national network to allow him to transparently access that facility as easily as any other facility.

7) ANSWER (cont.)

(b) The outstanding policy issue that I see coming quickly to the fore is that of intellectual property rights. The national centers, particularly if they attract considerable industrial funds, will have many types of people developing products which may be marketable. There will be individual faculty members at the university which runs the supercomputer, faculty members who are visiting that supercomputer center but are bound on intellectual property matters to their home institutions, the computer staff at the national center, the computer vendor partners who have donated large amounts of equipment to the national centers, and personnel from corporations who may be restricted by proprietary considerations at their institutions. Of course, because of Federal and State funding these conflicting rights must always be rectified with the needs of the government to obtain royalty-free licenses to developments coming out of these centers.

This rich research mix is a complex combination of players which will require that a great deal of attention be paid to matters of intellectual property rights. Again, I have been very impressed with the caliber of individuals that the national centers have working on these problems, and it should not be automatically supposed that a Federal solution is required. Nonetheless, the Congress should be aware that major conflicts will probably arise in specific cases at universities just as they have before in issues of computer software copyright and patents.

8) When you speak of a national network, are you referring to one U.S. network or a distributed network with gateways among numerous networks?

ANSWER:

I am concerned, as I believe Ken Wilson is, that the attempt to build one giant national network hooking each researcher at each university up together and with the national centers is almost surely doomed to failure by the very size of the undertaking. Since there are already in existence specialized networks, such as ARPANET for defense researchers or SPAN for NASA space project researchers, it may make sense to build the national research network in a more hierarchical fashion.

At least two of the centers already have regional networks hooking up consortia members on the east and west coasts. In addition, Cornell is helping the State of New York form a regional network of educational institutions. Illinois is considering forming a network of research groups from across the nation which add computational capability to our center, for instance in the areas of film animation or specialized software.

The first priority is then to provide a high capacity backbone network linking the four centers, each with their individual networks. The NSF policy is to establish 56kbaud links by early 1986. The funds need to be made available for those to be upgraded to megabaud lines as soon as possible. Furthermore, the four centers are going to become ARPANET nodes latter this year. NSF is joining DOD in adding many more universities to the ARPANET in 1986. Thus, by the end of next year, the top 30-40 research universities will have low speed connection with each other.

Congress should have hearings to determine the most appropriate manner in which to upgrade the network. This upgrading will consist of both adding more universities and of steadily increasing the carrying capacity of the network. To achieve this will require more funds in the NSF budget than are now contemplated. However, I strongly believe that the increased connectivity of our best researchers will more than pay back this investment in the years to come.

RECEIVED

AUG 27 1985

COMMITTEE ON SCIENCE
AND TECHNOLOGY

July, 1985

Response to "Questions for the Record"
from U.S. House of Representatives Committee on Science and Technology
by Kenneth G. Wilson

Q1. "Should NSF manage the new National Centers as a single resource or a collection of independent capabilities? Why?"

A. The National Centers should be managed as independent capabilities which are available on a single network. Each Center faces different problems in dealing with a different vendor or different technologies. Each Center has to be free to build its own distinct technological future and its own industrial partnerships; pooling the Centers would create an unmanageable conglomerate.

The great strength of the U.S. in supercomputing, as in many other areas, lies in the diversity of approaches that exist in supercomputing. A single resource would not adequately reflect this diversity.

Q2. "The NSF Bardon-Curtis Report projected the need for a NSF Advanced Scientific Computing program to be funded at $100 million per year. As the new NSF and DOE centers begin to train potential supercomputer users and new scientific problems requiring supercomputers arise, the demand for access will no doubt increase. What do you see as the long-range need for national supercomputer centers, mini-supercomputers centers, and local workstations?"

A. The future need for national supercomputer centers, mini-supercomputer centers, and local workstations will grow faster than the total government expenditures for university research and development.

The budget for these needs will be measured as a percentage of the total government R&D expenditures in universities, with this percentage increasing probably to the 8% level already reached at Los Alamos, as computing use in universities catches up with the level already experienced at Los Alamos.

Q3. "(a) How do you view the way in which NSF currently allocates supercomputer access time to researchers? (b) What impact will the current allocation of time to your Center have on its mission and the way in which its products and services are managed?"

A. It is too early to assess the impact of allocation procedures on the Cornell Center. We have not encountered problems yet.

Q4. "(a) Who establishes the policy for the allocation of supercomputer time at your Center? (b) Are research proposals for large segments of time, versus proposals for normal cycles of time, evaluated by the same criteria? (c) Are the criteria used by NSF to allocate time at the NSF National Centers compatible with needs of all disciplines, especially interdisciplinary research proposals?"

A. (a) NSF plans to allocate 60% of the time at Cornell. The Cornell Center for Theory and Simulation in Science and Engineering sets policy for the other 40%, with the help of an Allocations Committee that will include some nationwide membership.

(b) At Cornell, the treatment of large proposals will differ from small proposals only in the following respects:

i) their merits will be examined more carefully;

ii) fewer awards will be made; and

iii) attempts will be made to find experimental computing systems or other special arrangements with very high power to handle some of the largest requests, at Cornell or elsewhere.

(c) Cornell's Center for Theory and Simulation in Science and Engineering is itself interdisciplinary and will provide a basis for interdisciplinary proposals to receive allocations.

Q5. "You underscored the importance of supercomputers to industry as well as industry's contribution to the NSF National Centers. The major interest of the private sector will no doubt be technology development and application. Do you feel the time will come when the National Centers will be self-supporting, or will federal support always be necessary for experimental basic research?"

A. Federal support will always be important for very long-range basic research, for example, in high energy physics and many areas of astronomy. However, research which incidentally helps advance computer technology and its use is likely to be heavily cost-shared from industrial sources because of the immediate benefits of such technological advances.

Q6. "(a) What do you feel the federal role in support of supercomputer research, training and application should be? (b) What should NSF's role be? (c) In your view, are supercomputer policies and technology issues adequately coordinated among federal agencies?"

A. (a) Federal support should, as now, concentrate on advanced training and long-range research both for supercomputers themselves and their applications, while encouraging cost-sharing for projects with shorter-term benefits.

(b) The NSF should be the lead and coordinating agency for university support because it alone has the responsibility to see that the entire range of university research and advanced training activities get proper support. For

example, the NSF should coordinate the development of a computer network generalizing the Arpanet and should coordinate an effort to define the software for an advanced "scientist's workstation".

(c) There needs to be continuing pressure on the relevant government agencies to establish a coordinated networking effort.

Q7. "Clearly, other universities will want their own supercomputer. (a) Is there a hierarchy of centers needed or a variety of capabilities, i.e., by discipline, region, performance level, etc.? (b) What policy issues should be considered by federal agencies to provide for future national needs, coordination of federal and non-federal centers, and cost benefits such as economies of scale?"

A. (a) Planning for the future of supercomputing should be based on the idea of a network, with the placement of individual supercomputers on the network being a less important concern than the maintenance of the network itself. The network will contain a variety of supercomputers and other computers which will change rapidly due to technological advances. The top 40 research universities should all have major facilities of one kind or another on this network, just because of the importance of computing to every one of these universities.

(b) The basic policy issue is to establish the network. Coordination of the network among Federal and non-Federal agencies should establish the basis for coordination of facilities on the network.

Q8. "When you speak of a national network are you referring to one U.S. network or a distributed network with gateways among numerous networks?"

A. A distributed network with gateways among numerous subnetworks. My experience suggests that the only way both politically and technologically to build a national network is to build regional and local networks - such as a New York State network, a Big Ten network, a Princeton Consortium network, etc., each with around twenty to fifty nodes, and then link them together. Such a collection of networks could start with megabaud transmission rates (T1 lines) instead of the 56 kilobaud technology of the Arpanet, and then be upgraded further as needed. The T1 technology for networking seems to be ready now for regional networks and links between them, but not for a single national network of much larger scale.

LAWRENCE LIVERMORE LABORATORY

National MFE Computer Center

NMFECC

RECEIVED

JUL . 0 1985

COMMITTEE ON SCIENCE
AND TECHNOLOGY

July 26, 1985

Honorable Don Fuqua
U.S. House of Representatives
Committee on Science & Technology
Suite 2321 Rayburn House Office Bldg.
Washington, DC 20515

Dear Congressman Fuqua:

I enclose my responses to the questions for the record in your letter of June 25, 1985.

I wish to thank you again for inviting me to testify before the Subcommittees. I hope that the information presented will be useful in your deliberations.

Sincerely yours,

John Killeen

JOHN KILLEEN
Director
National MFE Computer Center

da

Enc.

University of California
P O Box 5509
Livermore, California 94550

Telephone (415) 422-4017
FTS 532-4017
Twx 910-386-8339 DOE LLL LVMR

MFE 85-241

QUESTIONS FOR THE RECORD

Dr. John Killeen

1. What do you see as the long-range need for national supercomputer centers, mini-supercomputer centers, and local workstations?

RESPONSE: It is gratifying to observe an interest at the Congressional level in the national need for adequate access to supercomputer power. I have been an active participant in several studies aimed at understanding and documenting the role of supercomputers in scientific research. It has been consistently found that current availability of modern supercomputer resources in the U.S. falls far short of the needs of researchers. It has also been projected that improvements in supercomputer technology will provide tremendous leverage in addressing forefront problems in contemporary and future science and technology. All computational models dealing with the frontiers of science and technology make simplifying assumptions about the laws of physics in order to keep the calculations from running too long. Improved supercomputers will enable researchers to develop more realistic models of physical phenomena. Such models have already been recognized as a cost effective alternative to the construction of expensive experimental devices which have been needed in the past to validate scientific theory.

We at the National MFE Computer Center have demonstrated that a supercomputer can be effectively shared by a widely dispersed user community. Although the unit cost of these machines is high, sharing the resource over a broad user base makes the cost per user very modest. Furthermore, we have shown that sharing the resource does not significantly interfere with optimum use of the machine by remote users. In my view, there could be no better way to improve the quality and impact of scientific inquiry in the U.S. than to provide the widest possible access to modern supercomputers. The development of computational methods needed to exploit supercomputers will be accelerated as their use becomes integrated into the scientific curricula offered at major universities. Equally as important, the firms now involved in the design and development of supercomputers will be induced to continue refinements and upgrades in their machines with confidence

that market potential exists for the product. Thus, world leadership in this field can be maintained by the U.S.

Since we have traditionally defined supercomputers as the largest and fastest machines in existence; the use of the term "mini-supercomputers" seems like a contradiction in terms. There has been a spin-off of smaller machines that have borrowed architectural and software features of supercomputers but not the large memory size or CPU speed. The idea of a center based in this technology would not be interesting to those who are striving to develop more realistic mathematical models. A center implies a substantial investment in communications to provide access. From my point of view this investment would yield a much greater return if the computer cited at the center is a supercomputer. It is more probable that a "mini-supercomputer" machine would function as a remote host at a network node and offer the supercomputer user a local resource that was software compatible with the supercomputer he was accustomed to using at the center. I envision the selection and use of these machines to be subject to local technical and economic decisions. A federal role in the promotion of access to machines of this class seems inappropriate to me. Similarly, I believe that competitive forces in the marketplace will produce an adequate mix of capable local workstations without federal inducements.

2. (a) How do you view the way in which DOE currently allocates supercomputer time to researchers? (b) What impact will the current allocation of time to your Center have on its mission and the way in which its products and services are managed?

RESPONSE: In my view, the allocation method used by DOE to distribute supercomputer resources is straightforward and efficient. The process of identifying and validating requirements begins with the submission of decision packages to DOE from all sites and contractors that require computing services from NMFECC. These decision packages are reviewed by DOE with respect to programmatic priorities. The available supercomputer resources are thus allocated in a manner

consistent with current DOE strategy for program mission accomplishment. NMFECC establishes computer "bank accounts" for all authorized sites and reports the use of resources to users and to DOE each month. (b) The NMFECC needs an allocation of computer time to conduct software development and research. Decision packages are submitted for DOE review and approval in the same manner as other users of the Center.

3. (a) Who establishes the policy for the allocation of supercomputer time at your Center? (b) Are research proposals for large segments of time, versus proposals for normal cycles of time, evaluated by the same criteria? (c) Are the criteria used by DOE to allocate time compatible with needs of all disciplines, especially interdisciplinary research proposals?

RESPONSE: The policy for the allocation of supercomputer time is established by DOE with input from the Computer Users Advisory Committee (CUAC) and the NMFECC management. (b) I believe that the predominant concern of DOE in evaluating decision packages is the degree to which the proposal aligns itself with prevailing DOE programmatic priorities. I am not aware of a bias toward proposals that request large vs small amounts of computing service. (c) Imbalance in access to supercomputers among disciplines has been sucessfully addressed through CUAC. For example, representatives of the fusion engineering community identified a variety of features that were needed to make the NMFECC more capable to run large engineering codes needed in the fusion program. These needs have been met without compromise to the interests of those who use the resources of the Center primarily for theoretical calculations.

4. (a) What do you feel the federal role in support of supercomputer research, training and application should be? (b) What should DOE's role be? (c) In your view, are supercomputer policies and technology issues adequately coordinated among federal agencies?

RESPONSE: (a) Because supercomputers are a powerful stimulus to

scientific advancement in many disciplines, the federal government should promote the broadest possible access. Prospects for improved supercomputer research, training and applications will all improve as a direct result of widespread availability. (b) DOE is uniquely qualified to make significant contributions to a national program to improve supercompter availability. DOE sponsored national laboratories have been at the leading edge of large scale computing since the 1950's. Typically the DOE laboratories have been the first to exploit new innovations in the large scale computing field by developing advanced operating systems, computational methods, and networking concepts. DOE's National Magnetic Fusion Energy Computer Center was the first to develop a national network for access to supercomputers ten years ago. The federal government should channel the supercomputer expertise that DOE has cultivated at the national laboratories into research that is vital to continued leadership for the U.S. in this strategic area. By this I am not suggesting duplicating or competing with industry, but rather establishing cooperative ventures. Because the supercomputer marketplace is small it is difficult for industry to profitably conduct the extensive research and development that is now required to press forward with promising new supercomputer architectures. With appropriate funding and coordination from the federal government, joint ventures between major laboratories and universities and U.S. computer manufacturers would hasten the development of supercomputer technology. (c) I have seen evidence of improved coordination among federal agencies during the past two years through the activities of inter-agency committees.

5. Clearly, other universities will want their own supercomputer. (a) Is there a hierarchy of centers needed or a variety of capabilities, ie.,by discipline, region, performance level, etc.? (b) What policy issues should be considered by federal agencies to provide for future national needs, coordination of federal and non-federal centers, and cost benefits such as economies of scale?

RESPONSE: The premise that universities will want their own supercompter may be correct,but I do not think it makes economic sense

to install a supercomputer for the exclusive use of a single site. Because of their high acquisition cost supercomputers should be fully utilized 24 hours each day of the year. This requires a large user base with a variety of computing styles. Having users in different time zones improves overall utilization of the resource. We should guard against the wasteful notion that a supercomputer is needed to make a statement of institutional prestige. (a) I see no reason to consider region or performance level factors in determining the establishment of national centers. Since a national center served by a national communications network tends to establish a "culture" of its own, it makes sense to create a community of researchers with similar or complementary disciplines. Collaboration and sharing are important aspect of a national center. The instant communications provided by the network is a powerful tool in itself. (b) The role of the federal agencies should be to advocate access to supercomputers as a means of stimulating scientific advancement in the U.S. To minimize cost, existing centers and networks should be expanded in scope until some evidence of diminishing returns has emerged. It is a major financial undertaking to establish the infra-structure needed to bring a supercomputer into operation. NMFECC has demonstrated that the same central facility, network, and talent pool needed to make one supercomputer accessible to a national user base, can be expanded to make several machines available. The cost of adding additional mainframes in established centers is only the acquisition cost of the new hardware and a modest expansion of the user support services (e.g., consulting) needed for an expanded user base. Experienced professionals in the supercomputer field are not easy to find and take years to develop. A pool of such talent represents a scarce resource that should be used to achieve the broadest possible benefit. Because of these cost factors, it is more effective to establish a few large centers of excellence rather than many small supercomputer centers.

The Florida State University
Tallahassee, Florida 32306-1047

Office of the Dean
Graduate Studies and Research

July 15, 1985

Honorable Don Fuqua, Chairman
Committee on Science and Technology
U.S. House of Representatives
Rayburn Office Bldg., Suite 2321
Washington, D.C. 20515

Dear Don:

Let me express my sincere appreciation, both personally and on behalf of the Florida State University, for having been given the opportunity to testify before the Subcommittees on Energy Development and Applications and Science, Research and Technology on June 10, 1985.

As requested in your letter of June 25, 1985 I have responded to the questions posed by the Members of the Subcommittees, that response being attached.

If I can be of further assistance, in this or any other matter, please call me at any time.

With best regards,

Bob

ROBERT M. JOHNSON
Dean, Graduate Studies & Research

RMJ/JWN:hs

RECEIVED
JUL 29 1985
COMMITTEE ON SCIENCE
AND TECHNOLOGY

RESPONSE TO QUESTIONS FOR THE RECORD

I. QUESTION: What do you see as the long-range need for national supercomputer centers, mini-supercomputers centers, and local workstations.

RESPONSE: The development of supercomputer centers, such as the Department of Energy center at FSU and the four NSF centers, will temporarily alleviate the urgent problem of providing a first level of access by scientists to supercomputers. It will not, by itself, solve the next level of the problem -- the need for more expert computer professionals who understand the architectures, the programming methodologies, and the relevant scientific disciplines.

This training need will best be met by dispersing the resources as wide as is economically feasible. Ideally this would mean a supercomputer at every university and college in the nation. Unfortunately, this is not feasible.

With that preamble, let me address the question. I think there will continue to exist a need for national supercomputer centers. In this sense, the word "supercomputer" must be taken in its original sense as being the largest, most powerful computer available at any given point in time. These supercomputer centers should always be at the cutting edge of the technology, offering the largest, most powerful machines available to scientists for attacking a wide spectrum of problems. Implicit in this answer is the urgent need for a national network which will enable scientists to access the systems.

I do not think that there is a need for national mini-supercomputer centers, although I think that federal assistance should be available to those educational institutions who are willing to assume a share of the burden. I believe that obtaining local involvement and commitments, sharing the cost of such centers with local government and private industry, strengthens local programs, and will ensure the success of those programs. The same is true of local workstations -- federal assistance should be available for assisting institutions who will put up their share.

II. QUESTION: (a) How do you view the way in which DOE currently allocates supercomputer access time to researchers.

RESPONSE: The present method of allocation appears to be reasonable in order to accomplish the goals.

(b) What impact will the current allocation of time to your Center have on its mission and the way in which its products and services are managed?

RESPONSE: The FSU Center assumes that 65% of the available

supercomputer time will be allocated by the Department of Energy via its Washington office, in accordance with already established guidelines. Therefore, such current allocation time will not adversely impact our mission nor management.

III. QUESTION: (a) Who establishes the policy for the allocation of supercomputer time at your Center?

RESPONSE: That 65% of the supercomputer time which is contractually obligated to the Department of Energy is allocated through the DOE's Washington office, as previously stated. The remaining 35% of the time is allocated by a Local Steering Committee, composed of faculty members who are qualified, by background and interest, to evaluate such requests for time on merit, and on need.

(b) Are research proposals for large segments of time, versus proposals for normal cycles of time, evaluated by the same criteria?

RESPONSE: While the same criteria - merit of the research -- is used, requests for very small amounts of time are usually for either training/familiarization projects, or for merely evaluating the feasibility of a project, and as such will be granted access much more readily than those large projects which are considered to be production research.

(c) Are the criteria used by DOE to allocate time compatible with needs of all disciplines, especially interdisciplinary research proposal?

RESPONSE: We feel that the criteria used by DOE could under certain circumstances exclude meritious research projects which do not necessarily fall under the umbrella of "energy related research." We cannot say that this has in fact happened, however, as we are not aware of any request being turned down. This may be due to requests only being submitted in accordance with DOE guidelines.

IV. QUESTION: (a) What do you feel the federal role in support of supercomputer research, training and application should be.

RESPONSE: The federal role in the area of research should be one of continuing to support worthwhile programs in the areas of basic research. Supercomputers are not an end in themselves, but merely a tool to enhance the capabilities of researchers. However, the power of available supercomputers is not sufficient to address many problems, and we can further anticipate that new areas of interest will be opened up as scientists gain access to these tools. The only known way of making supercomputers more powerful is to add processors, and the methodologies for properly utilizing a multiprocessor machine for a single problem (i.e., spreading the problem over several processors) are not yet developed. This is the most critical area of supercomputer research, and one which

must be strongly supported at the federal level.

In the area of training, I have previously addressed this issue, in my answer to the first question. Training will be enhanced by dispersing the resources over as wide an area as feasible, and ensuring that a communications network exists to allow many colleges and universities (and private industry) to link to these physical and intellectual centers.

In the area of applications, the federal role should be one of encouraging the wide distribution of public domain software, encouraging documentation and programming standards, and fostering interchange programs among the various institutions to ensure cross-fertilization and synergism.

(b) What should DOE's role be?

RESPONSE: DOE has long been the standard-bearer for the national laboratories, which represent a national resource of irreplaceable value. The roles of DOE and NSF are very different, and should not be diluted by duplication of effort for either. DOE should continue to support and nurture its national labs, but should also use these labs as centers of excellence for training, software development, and role models for developing supercomputer centers.

(c) In your view, are supercomputer policies and technology issues adequately coordinated among federal agencies.

RESPONSE: The establishment of FCCSET has strengthed the role of the agencies in this area, certainly, and it is not clear that in the area of supercomputers per se that there is inadequate coordination. However, in the area of networking, it appears that there is insufficient coordination. There is presently an uneasiness among the university community that various agencies will establish their own networks which do not communicate with other networks. While the current effort being spearheaded by NSF is laudable, it may be that the establishment of a separate effort charged with the implementation of a network for use by all agencies will be preferable in the long run.

V. QUESTION: Clearly other universities will want their own supercomputer. (a) Is there a hierarchy of centers needed, or a variety of capabilities, i.e., by discipline, region, performance level, etc.?

RESPONSE: This appears to align closely with the first question, and my response to that is also applicable to this one. Assuming that other universities do establish centers, either by full funding, partial funding, or entirely on their own, the issue is what sort of centers should be encouraged. I tend to think that such centers should be classified by discipline, and one of the criteria for selection would be the strength of a particular institution (or its willingness

to commit to establishing such strength) in a particular discipline which heavily uses supercomputers. (Examples are high-energy physics, oceanography, meteorology, quantum chemistry, etc.)

(b) What policy issues should be considered by federal agencies to provide for future national needs, coordination of federal and non-federal centers, and cost benefits such as economies of scale?

RESPONSE: I feel there will be a natural hierarchial progression imposed by financial constraints, willingness of vendors to support various areas, and the availability of over time of new architectures. This sort of variety should be encouraged by the policies developed by the federal agencies, as it will in the long run lead to the broadest possible training and development efforts. Policies which lead to a sharing of public-domain software should be encouraged. The use of centers of excellence, such as the national labs, for development of advanced software and training should be a shared effort. And most importantly, the development of a national data communications network, linking federal and non-federal centers, other institutions, and private industry, should be a national priority.

Mr. FUQUA. Our final wrap-up, cleanup speaker will be Dr. George Kozmetsky, director of the Institute for Constructive Capitalism at the University of Texas at Austin, and a very prime mover and shaker in this type of advancement. He has met with us on several occasions and has testified before us previously and has always been very enlightening in this particular subject matter, very positive and constructive statement.

Dr. Kozmetsky, we are very delighted to have you again.

[The biographical sketch of Dr. Kozmetsky follows:]

Dr. George Kozmetsky, Director of the Institute for Constructive Capitalism (IC^2) since its inception in 1977, is a Professor of Management and Computer Sciences and holds the J. Marion West Chair for Constructive Capitalism at The University of Texas at Austin. He also serves as Executive Associate for Economic Affairs to the Board of Regents, UT System.

From 1966–1982 Dr. Kozmetsky was Dean of the College and Graduate School of Business at UT-Austin. Prior to coming to the university, he was co-founder and Executive Vice President of Teledyne, Inc.

STATEMENT OF DR. GEORGE KOZMETSKY, DIRECTOR, INSTITUTE FOR CONSTRUCTIVE CAPITALISM [IC^2], UNIVERSITY OF TEXAS, AT AUSTIN, AUSTIN, TX

Dr. KOZMETSKY. Thank you, Mr. Chairman, members of the Science and Technology Committee, and the honorable Senator Gore.

The topic I would like to talk on is supercomputing and national policy, and the emphasis is going to be in maintaining our preeminence in an emerging industry. I am using industry in a very special sense, I am not talking about it either academically or in the traditional sense with which we have accepted it today.

First, we have not been able to define industry economically or from any theory I know today; it is very difficult to do.

Second, I think we have heard enough testimony to tell us this is not a traditional industry. So in a sense, supercomputers are an emerging industry. It is an industry that has been evolving over

the last 25 years and yet it is still in its infancy. Until 1980, it was predominantly an American development. Since 1980, it has become fiercely competitive and an international race for both scientific preeminence and economic preeminence.

We have concentrated mostly this morning on only half of the market or less, for supercomputers—namely research, industrial uses, Government uses and universities. We have not spent as much time as we will during this conference on the business and industrial market and perhaps the broader Government markets.

I think problems are just as eminent as the distinguished scientists from our great universities who have appeared before you the last 2 years and our Congressman from New York just in the last month. I know some of your companies outside of your district who are doing such great work in automating the factories are so frustrated in their laboratories because American management is sending the work over to the Far East for manufacture.

So, this is critical. In order to place our arms around this problem, I thought I would break my paper into three parts. Mr. Chairman, I do hope you will put into the record my complete written version.

Mr. FUQUA. We will make that part of the record.

Dr. KOZMETSKY. Thank you.

First, I would like to give some background for understanding the issues of industry perspectives, a light discussion of the issues and then finally a few conclusions.

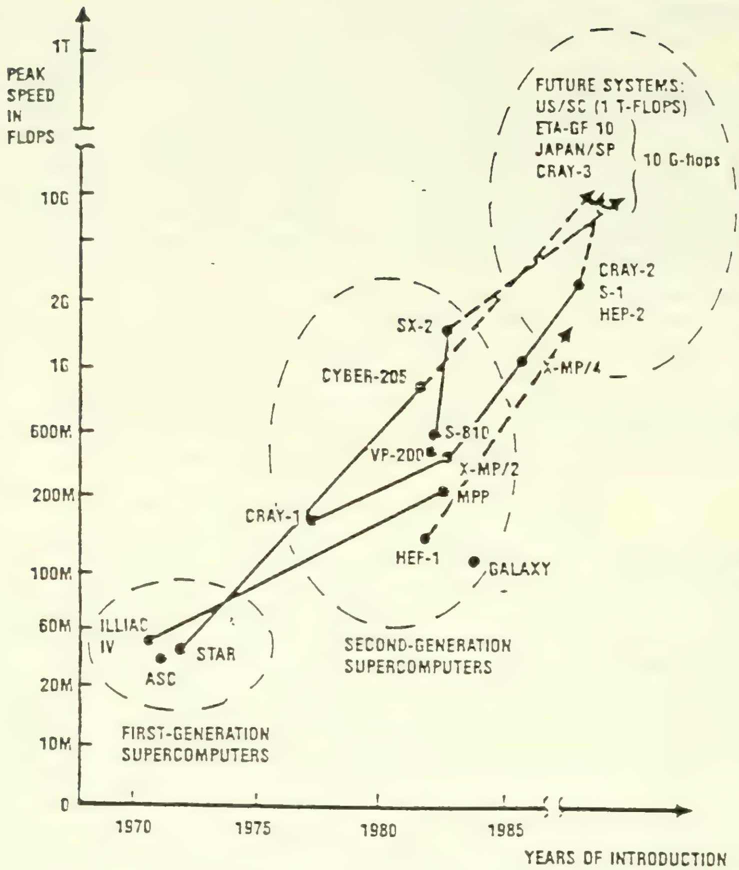
When it comes to getting a background, I am sure that all of our experts will agree with me in this room that the nature is complex, confusing and rapidly evolving. That is why we put an annotated bibliography together for this program, and I want to thank all of those who sponsored it. But for policy purposes, I think it is important to bring some structure to the area of supercomputers along with relevant data and information.

The following structure, I think, is relevant to understanding the emerging supercomputers. First, what I am going to call generation II computers; second, some current structure that I see emerging from it, and I thought I would bring you up to date on the predictions that I made November 16, 1983.

Generations	Time Period	Peak Speed in Flops
1. Beginnings	1959-60	1 - megaflops
2. First Generation	1970-74	10-100 megaflops
3. Second Generation	1975-1984	100-1000 megaflops
4. Third Generation	1985-1987	2 - 10 gigaflops
5. Fourth Generation	1991-	1-Trillion flops or more

SLIDE 1

Just to go over it very quickly, because you have heard so much about generations and some other words used to describe it. When Senator Gore asked his questions, I put it in the beginning, we were talking about one megaflop max speed at that time. First generations between 1970 and 1974, 10 to 100. There is the second generation when we start talking about the Cray 1 and Cray 2 for the experts, and the third generation is when we bragged about what America had. Then of course the CDC 205, the HEP–1 and some of the others. Then, I put in a fourth generation that generally covers the circumstances you have heard from the great, very great testimony from DARPA.



Development of modern supercomputers.

Source: Hwang, K. "Multiprocessor Supercomputers for Scientific/Engineering Applications." Computer, June 1985, p. 58.

Slide 2

So let us just simply see if what I see as the trends, from looking at these generations. Starting with June 1985, the first and third generations I already told you were delivered. Second, the economic and scientific global competition toward supercomputers if you read in the newspapers and magazines is in the third generation machines.

The U.S. Government, through DARPA at the Livermore Laboratory is developing the fastest and largest of the fourth-generation

machine. Their anticipated specification are two orders of magnitude faster than the Cyber 3, ETA and maybe supercomputer design specifications.

Fourth, all supercomputer-dominant companies, and I am going to use the word dominant to mean 50 percent or more, are considered emerging companies. None is a Fortune 500 or dominant in the large frame electronic data processing industry. None is vertically integrated as are the Japanese companies.

Fifth, the supercomputers are at a critical juncture. The second generation is being phased out as the third generation comes on stream. In the past such generation transition caused financial difficulties, and I will not say more on that.

Now, at the moment the supercomputer industry appears to be formless and highly unstructured. To put our arms around the state of the supercomputer industry, we must pull together various fragments that make up the industry. Let me tell you they are made up of universities, Federal Government agencies, it is made up by the State and it is made up by a whole series of collaborative institutions which the National Science Foundation calls institutional developments.

We at IC^2 in our research have consistently called it technology venture. It has university and Government relationships; it has university, Government and business relationships; it has State government, Federal Government, business. We have identified to least eight of these newer institutional developments and we can tell you in the NSF there are a number of gaps in institutions to really build an emerging industry.

I would like to show you table 1, (slide 3) which is a recent estimate we have made for the supercomputer industry. Much of it has been talked about here. We have several advantages at IC^2. We are not bound to be reelected, we are not bound to get State funds to support us and we can look 10 years ahead. So these are strictly estimates of IC^2, and the first thing you notice is that there is over $8 billion. Now that total funding for supercomputer research is almost as much as NASA's projections for a manned space station. It is a significant amount. The estimate of funding which you saw flows from three sources.

TABLE 1

Estimated Direct and Indirect Support
For Supercomuter Developments
Next Ten Years
(In millions of dollars)

1.	Federal Agencies	
	A. DOD – Strategic Computing Survivability Program	$ 1,000
	– Strategic Defense Initiatives, Robotics and Artificial Intelligence	1,200
	B. NSF – Supercomputer University Centers	200
	– Communication Network	5
	C. DOE	200
	D. NASA – Space Station, Automation, Robotics and Artificial Intelligence	600
	E. NIH – Medical Information Systems, Biotechnology Knowledge Bases	100
	F. Other	45
		$ 3,500
2.	Supercomputer Companies	
	A. Primary Manufacturers	1,500
	B. Secondary Firms and New Start-Ups	1,000
		$ 2,500
3.	Cooperative R&D	
	A. MCC – Consortium	1,000
	B. Semiconductor Research Corporation	100
	C. Stanford University Center for Integrated Systems	300
	D. Massachusetts Microelectronic Center	100
	E. California Microelectronics Innovation & Computer Science Program	100
	F. Minnesota Microelectronics & Information Science Center	100
	G. North Carolina – Microelectronic Center	100
	H. Florida State Supercomputer Computations Research Institute	100
	I. NSF University Supercomputers Centers – Matches	300
	J. Washington-VLSI Technology Consortium	15
	K. Indiana Computer Integrated Design, Manufacturing & Automation Center	20
	L. Other	50
		$ 2,285
	TOTAL	$ 8,135

SLIDE 3

Now we projected Federal Government funding, as you can see, at $3.4 billion. All I am going to do for this part of the oral testimony is tell you the dominant portion is coming from DOD. They have enough experts here to criticize my projections and they might be happy or unhappy, but that is all right. We learn at IC^2 and improve it for the next testimony. The private sector is estimated to provide about $2.5 billion, and as you can see, the dominant is to come from the three primary small company manufacturers.

That means the market must continue, the prices must stay up. Comparative R&D support is composed of State government, private corporations, universities. The dominant portion will be from private company research programs and that is mainly from MCC.

Now the supercomputer companies are directly involved with the Japanese computer race in the marketplace. Both U.S. responses to the Japanese fifth generation computer and the supercomputer scientific and engineering race from an R&D perspective are taken up by private companies forming MCC and DARPA. To the best of our knowledge the current three supercomputer dominant companies are not directly involved.

As an infant industry, the supercomputer industry is in an evolutionary stage. Since my last testimony before this committee on November 16, 1983, the industry has taken on a different nature. The industry has become more diversified by broadening its base with slower machines for computational purposes. It now can be described as follows: supercomputer manufacturers and super-minicumputer manufacturers, add-on main frames and the flexible machines and software—I guess we left that off the slide (slide 4).

1. Supercomputer manufacturers;

2. Superminicomputers manufacturers;

3. Add-on to mainframes; and

4. Software, graphics, and other peripherals.

SLIDE 4

Now I think it is important for us to get a benchmark that there are about 106 systems that have been delivered to date (slide 5). I did not say central processors, there are more than that, there may be 120. As of June 1, we think 106 systems have been delivered. Cray is dominant, followed very closely by ETA and Denelcor and Amdahl has yet to deliver any. But we can tell you all four manufacturers in the United States, and I have included Amdahl because they were involved with the Fujitsu design of supercomputer, but they all have substantial backlog for this second and coming third generation into the 1987 and 1988 timeframe.

Manufacturers	As of June 1985 Estimated Number of Systems Delivered
Cray Research, Inc.	61
ETA Systems, Inc.	39*
Denelcor	6
Amdahl	0
	106

* Includes Cyber 205 only which is manufactured by CDC

Slide 5

Now the applications for supercomputers are still those that I presented in my testimony on November 5 and we will just flash this up. What you have seen there, I hope it is familiar to all of you and I will not bore you any more on that (slide 6).

TABLE 3

Effects of Preeminence in Supercomputer Technology on the Computer/Communications Industry

Drivers	Area Affected

Industries Developed and Expanded by Super Computer Related Technologies	Robotics/Industrial Automation
	Computer-aided Engineering
	  Computer-aided Design
	  Computer-aided Manufacturing
	  Computer-aided Testing
	  Computer-aided Quality Control
	Biotechnology
	Office Automation
	International Financial Services
	Service Industry
	Computer Industry
	Semiconductor Industry

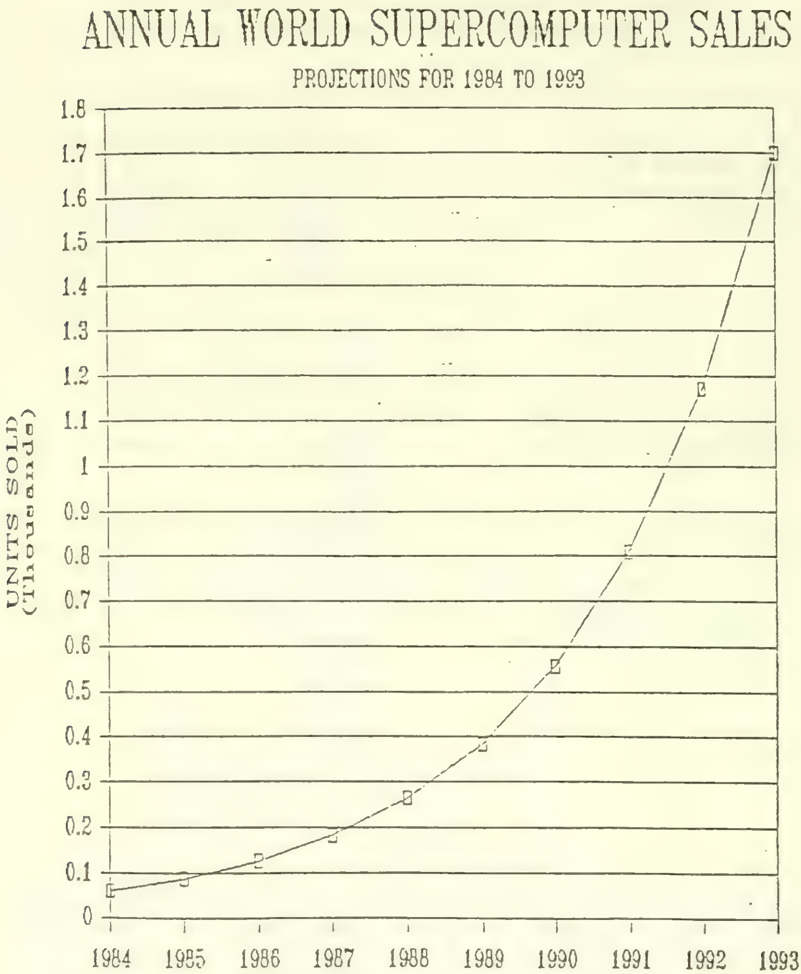
Revitalization of Basic Industries	Energy
	  Petroleum
	  Nuclear
	Steel
	Automobile
	Textile
	Chemicals

Institutions Receiving Significant Productivity Boosts	Government Defense Education Research
Communications Technology Enhancements	Subscribers Equipment Telephone Co. Equipment Overseas Carriers Satellite Carriers

SLIDE 6

Now the current market trends for annual worldwide supercomputer sales indicate the sales on units sold are on track according to that forecast we made for you in November 1983. I thought you would like to see how we came out dollarwise, so on chart B (slide 7) we are going to show you the early projections, we took our conservative ones only and you will notice that we missed by $200 million, and that is because we were not smart enough to think of leasing machines.

Chart B



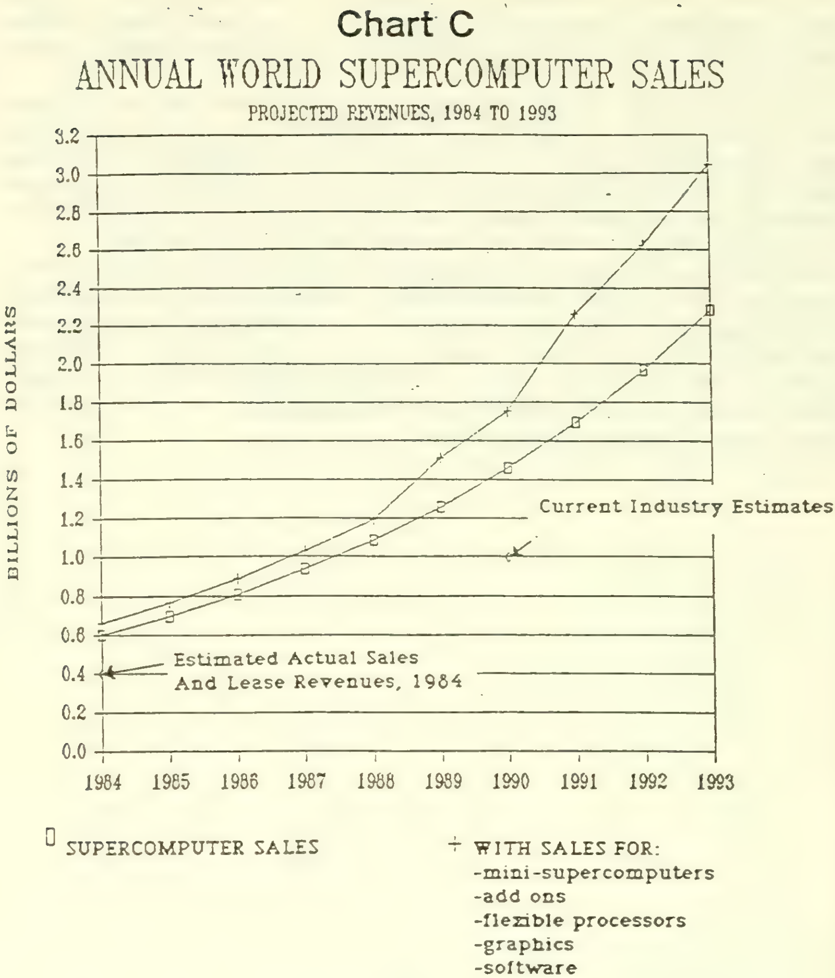
Source. IC² Institute, The University of Texas at Austin.

SLIDE 7

In 1983, I did not think any buyer of a supercomputer would lease a machine. If you took the leasing in, I am sure I am at worse maybe 20 percent off on the projection. I think I am slightly closer, in fact.

Now the trends in the industry can be summarized as follows, and I think first is an increase in demand for supercomputers that was unforeseen as short as 18 months ago. Demands are arising from the Federal Government needs, expanding industrial and business markets both domestic and foreign, and universities. Government is getting to be a decided less factor in the market.

All four supercomputers companies, as I told you, have robust backlogs. The smaller supercomputer manufacturers are finding a developing market for their products in educational research labs and we predict that their sales are going to go up. I have now modified my prediction because what is happening, as you have been hearing in all the testimony this morning, we are beginning to get mini supercomputers (slide 8).



Source: IC[2] Institute, The University of Texas at Austin.

SLIDE 8

Some of us in education are going to argue slightly about teaching with personal computers. I always take about 20 million students multiplied by $1,000 and say, how can we afford $20 billion of personal computers?

I think we all mean the same thing, we want nice, smart, simple terminals that ought to be manufactured in volume and made reasonably cheap in the few hundred dollar range, so that we can tie in with these bigger computers.

So I have added in mini-supers and flexible machines, some of the newer things that will be discussed today, and the software. And I think that for the committee to get a ballpark figure of why it is still an emerging industry in 1993, main frame computers in 1984 sold $17 billion. I do not think I have more than $3 billion up there in my perhaps unoptimistic forecast.

Now let me get to discussing policy. How will supercomputer R&D continue to be performed and funded in the future? What are the changing patterns? For the initial supercomputers and the first general of supercomputers, R&D was funded by industry and Government.

The first customers for a deliverable second generation supercomputer were Government labs.

But the third generation of supercomputers, R&D is being funded by industry. This is the generation about which I have already said that the Japanese have announced that they want the competition.

The first known customers of U.S. deliverable third generation machines are Government labs. As a matter of fact, Cray 2 is the beginning and it has already been delivered, and university supercomputer centers funded jointly by the Federal Government through NSF and the State governments and computer manufacturers and perhaps other businesses.

The fourth generation, R&D is funded by the Federal Government. I have not found any private funding in the companies I have visited in the fourth generation.

As I said earlier, starting with the fourth generation, U.S. specifications are substantially ahead of the current Japanese 1990 goals and are in some respects a way of meeting the Japanese. To the best of my knowledge, none of the four supercomputer manufacturers in the second and third generation are actively involved in the fourth generation.

Let us see where we can get the funding to do R&D. What we are getting in our institutional developments or technology ventures, if you like, continued R&D companies without Government funds. You get a ballpark third generation—second generation costs on the average about $90 million with no provisions for the technology you need to develop supercomputers, probably went about $120 to $150 million in capital.

Second, joint venture of companies, including foreign firms and licensing of technology. Third, participate in cooperative research and nonprofit consortiums such as MCC for staying ahead in technology. Next, enter into R&D contracts with Federal agencies for a specific supercomputer or enabling technology for developing supercomputers.

Next, sell equity to other corporations but also license for their purposes the technology being developed by supercomputer hard-

ware itself, and that generally means the giant corporations of America.

Next, utilize R&D partnerships and other traditional capital venture processes and I can tell you when you look at $4 billion going out in 1984 for those traditional things and looking at $106 billion for mergers and acquisitions, $18 billion for leveraged buyouts, forget it.

Next, have Government laboratories or universities buy or lease early models and next generation supercomputers to help finance supercomputer developments; manufacturers as well as use them as sites and thus securing enough operational time for buying or for leasing than in some cases even the Government laboratories would be able to purchase.

And finally, establish a collaborative link with research performed by Government laboratories and universities. These links can be utilized for relevant intellectual properties funded by others such as the Federal Government or State government, or directly funded by gifts or business donations.

Other collaborative linkage would be to sponsor university research for commercialization developments and then retain all of the commercial rights, and I am sure there are way over 50 supercomputer developments going on in our various universities in this country already.

Pretty soon, every good computer scientist will have his own desk drawer plan for little old supercomputers way beyond the fourth generation.

Now, what kinds of complex applications are in the nature of developments on supercomputers? I do not think we have been very frank with this committee. I have not seen anything done by supercomputers other than the problems that I worked on in 1950.

They were all for national security; complex, large, and we did not have either the scientific knowledge, mathematics, computer hardware of software to solve them, and we have been working on them for the last 30 years.

Now we have tried to tell the committee and I think it has been done eloquently, it says we need to be working on three types of large scale problems beyond the capacity of the fourth generation.

First, I will not bother to detail, are large scientific research projects. That is where you get the Nobel Prizes, that is where you get the prize for a nation.

Perhaps I was the only one listening, but we need a second class of problems for advanced flexible manufacturing and processing computerized systems.

There is nothing in existence to handle those. We heard about the whole automobile design. If we are going to be ahead in this Nation, that is a big problem to work on and there is no computer, mathematics, science or other things.

I think it is extremely important and that we have an awful lot of things to do already and that is advanced security needs and applications. There are still continuing supercomputer—you can see the details on my slides and you can see they are all very large markets.

Let us go to the third one. Cooperative research or technology ventures, reduce the development time and maintain the United States competitive position in the global supercomputer market.

Recent institutional developments for cooperative research and development includes supercomputer applicability initiated primarily for economic growth, job creation, and diversification, not for science.

These cooperative efforts include industrial consortiums, software consortiums, university consortiums and State industry cooperative programs with the universities. Government labs have already been directed by this committee and others that they are going to actively pursue collaborative efforts to transfer and defuse their technological effects, yet it is highly unstructured and how policies are formulated will go a long way towards making and securing our Nation's and our individual future as well as the future of the emerging supercomputer industry.

The Japanese seldom have used academia in their actual problem solving. They more often utilize business and Government laboratory personnel. They also establish timeframes for projects much like our sunshine laws. They also accomplish reviews when the researchers under their sunset laws return to their firms or Government laboratories.

Now this is not necessarily what the United States requires. Our social and work values are different. We do not utilize lifetime work principles in employment and neither our firms or Government laboratories—I was long enough in the university life to know.

The newer supercomputer research centers can modify and promulgate new approaches to better enable their research to be more rapidly and efficiently incorporated. With this approach, past time spans of 15 to 30 years in technology and the economic products and services can be successful. The utilization of an intellectual property right can competitively produce research results.

The fourth issue I would like to discuss, how can the results of cooperative research on supercomputer technology be transferred? With Government labs and university supercomputer centers as a primary focus, we can illustrate the decision.

Other cooperative research, institutional developments that involve local, Federal and State governments, businesses and academic institutions and collaborative efforts.

The national laboratories where centers are established produce among other things national technological—how to distribute this equally or fairly among the regions of our country is a newer issue.

In the past, since 1980, various States have been concerned on job creation and the utilization of higher technology or 150 initiatives have been picked over the past 5 years by the States.

The IC^2 Institute at the University of Texas has conducted a study to determine where the developments are taking place. For these purposes, the major drivers I used were Federal tax, probably \$3 billion. And selected R&D expenses in the selected companies and those who have 1 percent of their sales in R&D, \$8 million. That is \$41 billion.

Now the States by ranking you are seeing, (slide 9), as you can see the dominant States, there are 17 of them and the District of

Columbia, that we see. Four of them are what I call the first tier; California, New York, New Jersey, and Massachusetts.

TABLE 4

Innovation State Rankings

State	Dominant Federal R&D Obligations[1]	Dominant Selected Company R&D Expenses[2]	Traditional Venture Capital[3]
California	1	3	1
New York	3	1	2
New Jersey	8	3	10
Massachusetts	4	5	3
Maryland	2		6
Texas	7		8
Illinois			5
Virginia	5		
Florida	9		
Nex Mexico	6		
District of Columbia	10		
Michigan		2	
Pennsylvania		5	
Connecticut			4
Colorado			7
Minnesota			9
Delaware		5	
Ohio		5	

[1]For fiscal year 1983

[2]1983-84

[3]1984

Source: IC^2 Institute, The University of Texas at Austin.

SLIDE 9

The second tier is Maryland and Texas. For the benefit of the audience and especially this distinguished committee, you will find California, New York, Texas, Illinois, Florida, Pennsylvania, and New Jersey fairly dominant.

So while transformation has taken place in terms of innovation, it has become evident that science and technology naturally evolved in commercialization.

It does not provide employment opportunities, mitigate layoffs, maintain a strong basic research posture, maintain a strong competitive position, and present opportunities across the board.

And, furthermore, the mechanism of allocating resource needs to focus more on flexibility than on efficiency and effectiveness.

Now the final issue, how can supercomputer advances be transferred and defused to small- and medium-sized companies? Supercomputers as well as mini supercomputers in many respects are expensive and beyond the reach of many small- and medium-sized firms.

Such sized firms have already been noted for their innovative abilities and as a major source for employment growth. As supercomputer development promulgates, they will be among the first to utilize the advances especially in design, quality control, et cetera.

In some sense, if they could be placed in a position to utilize these results, they could help stem the flow of jobs outside the United States and at the same time increase other employment opportunities through new product developments and increased productivity.

Currently there has been little concern or organized effort to examine this need, and it has come up already in the testimony today.

This aspect could be achieved by incorporating some measure that utilizes the university supercomputer center network system.

This will help in developing a standardization of supercomputer hardware, peripherals, station terminals, graphics, software, networks; and I think the networks for industry are going to be different than the networks for the research labs, Government labs and for universities. And I think they have got to be different for business and Government, but the important thing that if it is done properly, especially on all the networking you have heard, it would gain us 3 to 8 years advance over both electronic data processing innovations, office automation and put us in the forefront in the United States over the world.

So, in conclusion, it is fairly clear that the supercomputer industry is in its infancy, it is in a promising period of growth, it is at a critical juncture, its product lines are becoming diversified as we said, in terms of super minicomputers and others, software and peripherals.

Today, it is an industry principally dominated by small, nondiversified supercomputer companies that have managed to meet the Japanese challenge. They are successfully dominating the global supercomputer market.

A long-term future for the supercomputer industry is still promising, it depends on how the user industry markets are developed. These markets are not necessarily replacement or substitution for current obsolete equipment, methods and services.

The markets are here and how we meet them with vision will give rise to new American dreams.

This committee can do much in formulating appropriate science and technology policy to assure that American small business and innovative enterprises, as well as our universities, our research labs and business and Government are assured of the availability of supercomputer technology at an earlier date than our traditional process of commercializing research, than was made possible previously.

Thank you for having me here, Mr. Chairman.

[The prepared statement of Dr. Kozmetsky follows:]

SUPERCOMPUTERS AND NATIONAL POLICY: MAINTAINING U.S. PREEMINENCE IN AN EMERGING INDUSTRY

by

George Kozmetsky

Supercomputers is an emerging industry. It is an industry that has been evolving over the last 25 years. Yet, it is still in its infancy. Until 1980, it was predominantly a U.S. development. Since 1980, it has become a fiercely competitive international race for scientific and economic preeminence.

A number of key policy-related issues must be addressed to ensure a robust U.S. supercomputer industry and to enhance its impacts on other industries and society. Among these issues are:

1. How will supercomputer R&D continue to be performed and funded in the future? What are the changing patterns?

2. What kinds of complex applications and needs will drive the advanced developments of supercomputers?

3. Can cooperative research reduce the time required for supercomputer development and maintain the U.S. competitive position in the global marketplace?

4. How can the results of cooperative research in supercomputer technology be transferred and diffused regionally and institutionally?

5. How can supercomputer advances be transferred and diffused to small and medium sized companies?

Acknowledgments

There are a number of friends and colleagues whose inputs and critiques have been especially insightful and helpful in the development of this paper.

I would like to express my appreciation to Dr. Hans Mark, Chancellor of the University of Texas System. From The University of Texas at Austin, I would like to acknowledge Gerhard J. Fonken, Vice President for Academic Affairs and Research, and Dr. James C. Browne, Chairman of the Computer Science Department. From the IC^2 Institute of The University of Texas at Austin, I would like to acknowledge the constructive suggestions of the following research fellows: Dr. Eugene B. Konecci, Dr. Ray W. Smilor, Professor W.W. Cooper, Dr. J.R. Kirkland, Michael D. Gill, Jr., and Dr. Stewart Nozette. I wish to thank Linda Teague for her invaluable assistance in preparing this manuscript and Patricia Roe and Ophelia Mallari for their research assistance and other support. Dean John Rouse and Dr. Ted Sparr of The University of Texas at Arlington and its Advanced Robotics Research Institute were especially helpful in providing insight for advanced supercomputer applications.

My appreciation also includes other academic colleagues: Professor C.V. Ramomoorthy of the University of California, Berkley;

Professor Edward Glaser, Former Director, Jennings Computation Center, Case-Western Institute; Professor Robert M. Fano, Department of Electrical Engineering, Massachusetts Institute of Technology; Frank Davidson, Systems Dynamics Group, Massachusetts Institute of Technology; Professor Herbert Simon, Carnegie-Mellon University; Dr. Jesse Poore, Georgia Institute of Technology; and Dean Robert Johnson, Florida State University.

The remainder of this paper is in three parts:

Part I - Background for Understanding the Issues

Part II - Discussion of the Issues

Part III - Conclusions

Part I -- Background for Understanding the Issues

The nature of supercomputers is complex, confusing, and rapidly evolving. An "Annotated Bibliography of Literature on Supercomputers," compiled by Dr. James Browne and John Feo of the Department of Computer Sciences, and Patricia Roe of IC^2 Institute, all at The University of Texas at Austin, shows that most of the literature is in terms of applications and architecture. The few articles dealing with business and marketing aspects of supercomputers are either company-distributed literature or generalized interviews with selected individuals. There is yet to be developed a body of literature which structures both the academic and professional fields of supercomputers, focuses on policy and social implications, and examines its business and industrial applications. This is another means of confirming that the supercomputer industry is in its infancy.

For policy purposes, it is important to bring some structure to the area of supercomputers along with relevant data and information. The following structure is relevant in understanding the emergence of the supercomputer industry.

A. Generations of U.S. supercomputers;

B. Current structure of the emerging U.S. supercomputer industry;

C. Current markets, applications and trends.

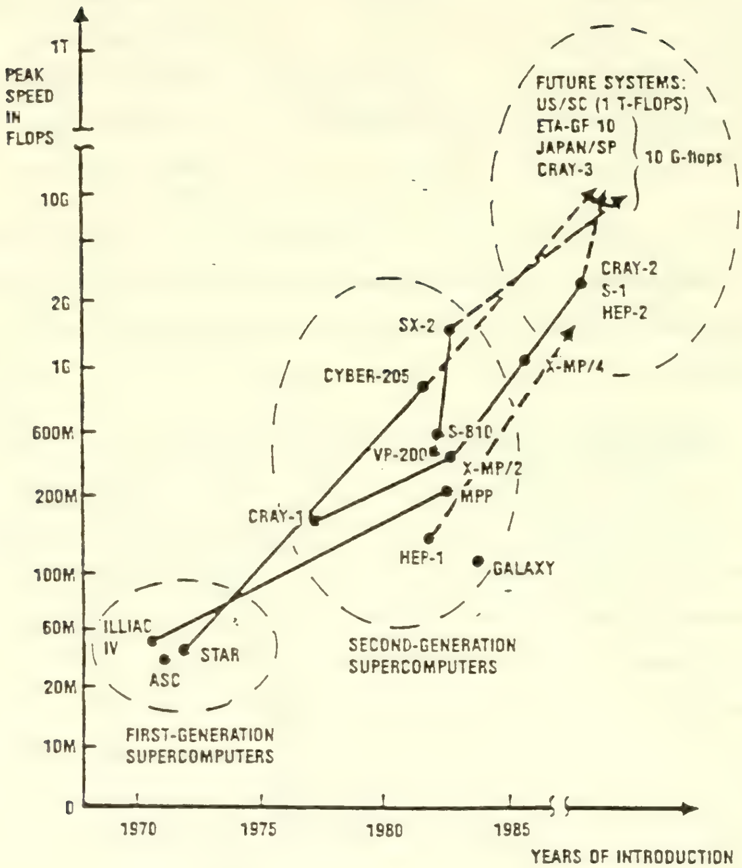
A. Generations of U.S. Supercomputers. There is a tendency by all technical persons to discuss supercomputers by generic names or numbers, by performance characteristics, by component contents, or by company names. When an industry is emerging, this adds to the confusion of the public's understanding of the industry's developments. To help clarify the picture, we have structured the generations of supercomputers based on the works of Professor Kai Hwang,[1] Sidney Fernback,[2] and Lloyd Thorndyke.[3] We have taken the liberty of assimilating their data to fit into the following categories of generations:

Generations	Time Period	Peak Speed in Flops
1. Beginnings	1959-60	1 - megaflops
2. First Generation	1970-74	10-100 megaflops
3. Second Generation	1975-1984	100-1000 megaflops
4. Third Generation	1985-1987	2 - 10 gigaflops
5. Fourth Generation	1991-	1-Trillion flops or more

Chart A shows the time of introduction and peak speeds for each of the manufacturers for the first through the fourth generation of machines.

Beginnings -- Supercomputers started in the late 1950's as R&D projects for scientific/engineering computations in the 1 megaflop range for the Livermore and Los Alamos National Laboratories. IBM and UNIVAC were the prime contractors to deliver what were called the STRETCH and LARC computers. These first machines were "shared risk" developments with the government ensuring the purchase of more than one machine from each vendor. These programs underwent problems that so many other early projects experienced. They underestimated the technological complexity of the project. As a consequence their deliveries were late; they incurred substantial financial overruns; and software was minimal. The early commercial supercomputer markets were captured by Control Data Corporation (CDC) with their CDC 1604. IBM and UNIVAC incorporated their gained technology in mainframes for the business and industrial fields.

Chart A



Development of modern supercomputers.

Source: Hwang, K. "Multiprocessor Supercomputers for Scientific/Engineering Applications." Computer, June 1985, p. 58.

First Generation -- In the early 1970s, government labs were instrumental in starting the first generation of supercomputers to perform as "multiprocessors" and "vector processors." Three manufacturers responded to their requirements: Burroughs, Texas Instruments, and Control Data Corporation. Burrough's ILLIAC IV was their only experimental model. It was built at a reported cost of $31 million, and NASA's Ames Research Center was able to achieve the performance levels for which this machine was designed. After more research, Burroughs decided to drop out of the supercomputer race.

Texas Instruments decided to venture into the supercomputer area by building their ASC (Advanced Scientific Computer). They built seven systems and delivered six. One went to the Geophysical Fluid Dynamics Laboratory, Princeton, NJ. One went to the Naval Research Laboratory in Washington, D.C. A third went to the Ballistic Missile Defense Laboratory in Huntsville, Alabama. The others were for internal use. TI is apparently no longer in the supercomputer field.

CDC developed STAR 100. Four of these were built. Two were delivered to the Lawrence Livermore National Laboratory and one was delivered to NASA/Langley. CDC kept one for internal use. While they discovered the reasons for the STAR's relative ineffectiveness for scalar calculations, they stayed in the supercomputer race with the CDC 6600 and 7600. In effect, they produced the progenitor for the second generation.

Second Generation -- Seymour Cray was the chief architect for CDC's supercomputers. He left CDC in 1972 to form Cray Research,

Inc., where he designed and built Cray-1, the first vector/scalar machine. As with the first generation, the first markets were the government labs. However, Cray-1 had a difficult time in selling its first computer to a government lab because of government regulations that required the machine be proven first. A unique six-month free trial arrangement with Los Alamos National Laboratory resulted in the purchase of the Cray-1; thus, the company overcame some of the regulatory obstacles.

In the meantime, a number of Cyber models were developed for the second generation. CDC renewed its interest in supercomputers in the 1970s. It improved its STAR 100 components and added a scalar arithmetic unit. The early models were called CYBER 203 and were delivered to NASA/Langley, the Navy's Fleet Numeric Oceanographic Center in Monterey, California and the CDC Data Center in Minneapolis. Enhancements to CYBER 203 were made, and the improved machine became CYBER 205. In 1983, CDC decided to spin off its supercomputer effort because they believed that a small company could do a better job of developing a supercomputer than a large company. ETA Systems, Inc. is the spinoff. ETA later acquired the rights to market the CYBER 205 from CDC.

Denelcor became the third U.S. company to enter the second generation of the supercomputer race. Their machine is a non-vector, multiprocessor. It is called a Heterogeneous Element Processor or HEP.

Three Japanese firms entered the supercomputer race to try to meet the needs of their Japanese customers. The first Japanese company to announce entering the race was Fujitsu. Their second generation machines are called VP 100 and VP 200 with peak speeds of 250 and 500 megaflops respectively. Hitachi was the next with their S810/10 and S810/20 with 315 and 630 megaflops respectively. NEC was the last Japanese company to enter the race. They announced their SX-2 to be delivered in 1985 at 1.3 gigaflops.

Fujitsu currently owns 49 percent of Amdahl Corporation. They have announced that Amdahl will market the Fujitsu VP 100 and 200 vector computers to operate with IBM's System/370 software.

Third Generation -- The third generation is in many respects a future system, meaning that it is still under development for later delivery. The companies in the second generation are all involved in the third generation. A brief listing of their machine titles and anticipated peak performance characteristics follows:

Company Name	ANTICIPATED		
	Machine Designation	Speeds (gigaflops)	Delivery Dates
Cray Research, Inc.	Cray-2 Cray-3	2 - 3 10	1985 1985-86
ETA Systems	ETA-10	10	1986
Denelcor	HP-2	2 - 3	1987

There is underway fourth generation supercomputer programs. One of these is DARPA's one trillion flop machine being developed under the Strategic Computing Survivability Program. Currently, it involves $1 billion to be spent between 1984 - 1991 in four areas: extension of artificial intelligence, multiprocessor architecture, advances in VLSI, and rapid turnaround and fabrication of integrated circuits. DARPA's goals are substantially above Japan's MITI's national supercomputer program for a 10 gigaflop supercomputer in 1990.

Livermore Labs, under the auspices of the U.S. Navy, is developing the S-1 multiprocessor project with speeds up to 90-1500 megaflops. At the high end of S-1 speed, it would be the fastest supercomputer currently being developed.

There are some important implications as well as trends that are evident from this brief review of the supercomputer generations; namely:

1. Starting in June 1985, the first of the third generation machines, Cray-2, will be delivered. Others will be delivered in the 1986 and 1987 timeframe.

2. The economic and scientific global competitive race for supercomputers is in the third generation of machines.

3. The U.S. government through DARPA and Livermore Laboratory is developing the fastest and largest of the third generation machines. Their anticipated specifications are two orders of magnitude faster than Cyber 3, ETA-10 or MITI's supercomputer design specifications.

4. All dominant U.S. supercomputer companies can be considered emerging companies. None is in the Fortune 500 or dominant in the large mainframe electronic data processing industry. None is vertically integrated as are the Japanese companies.

5. Supercomputers are at a critical junction. The second generation is being phased out as the third generation comes on stream. In the past, such generational transitions caused financial difficulties. The problems are likely to occur again because of manufacturing start-up costs for the third generation and shut-down expenses connected with the second generation.

B. Structure of the U.S. Supercomputer Industry. There is no question that the U.S. government has been and still is the major

investigator for supercomputer research and development and a key user of computers. One can question why the supercomputer industry has taken so long to develop. The answer is that to date larger scale computing markets were outside the purview of the supercomputer. When the early supercomptuers were developed for the government, their computing power and other abilities could be marketed or surpassed by large mainframe computers which manufacturers could develop to meet business and industry applications. As a result, mainframes became the core of the U.S. computing industry. Furthermore, they were very profitable. The rapid pace of developments of large scale processing resulted in newer generations of machines for data processing with large data bases and an ever-growing market. Today the mainframe computer business to over $17 billion. With over 106 supercomputers systems sold and delivered to date, their markets are still approximately $500-600 million a year.

At the moment, the supercomputer industry appears to be formless and highly unstructured. To put our arms around the state of the supercomputer industry, we must pull together various fragments that make up the industry. This includes R&D, supercomputer manufacturers, support infrastructure, software, and communication networks.

Research and Development -- Research and development for supercomputers is more indirect than directed. Over 1/3 of the directed research and development for the supercomputers during the next decade will be funded by the government. The DARPA-Strategic Computing

Survivabilty Program research and development, while being conducted in four different areas, still has a 1-T flops supercomputer to be developed as an end objective. The Livermore S-1 also has a specified supercomputer as a major part of its program. The other major supercomputer R&D programs are by the major manufacturers who are developing their supercomputers for the third generation, namely Cray, ETA, and Denelcor.

On the other hand, there is a larger amount of indirect R&D support that can be utilized for supercomputer R&D. These efforts are highly diffused and fragmented. Table 1 shows the IC^2 Institute's recent estimate that over $8 billion will be invested in both direct and indirect R&D for supercomputers.

The total funding for supercomputer and related research is almost as much as NASA's projection for the manned space station. It is a significant amount. The estimated funding is coming from three major sectors. We have projected federal government funding at approximately $3.4 billion with the dominant portion coming from the Department of Defense (DOD). The private sector is estimated to provide direct support of over $2.5 billion. The dominant portion is forecasted to come from the three primary manufacturers. Cooperative R&D support is composed of state governments, private corporations, and universities. The dominant portion would be private corporate research program support for the Microelectronics and Computer Technology Corporation (MCC).

TABLE 1

Estimated Direct and Indirect Support
For Supercomuter Developments
Next Ten Years
(In millions of dollars)

1.	Federal Agencies	
	A. DOD - Strategic Computing Survivability Program	$ 1,000
	- Strategic Defense Initiatives, Robotics and Artificial Intelligence	1,200
	B. NSF - Supercomputer University Centers	200
	- Communication Network	5
	C. DOE	200
	D. NASA - Space Station, Automation, Robotics and Artificial Intelligence	600
	E. NIH - Medical Information Systems, Biotechnology Knowledge Bases	100
	F. Other	45
		$ 3,500
2.	Supercomputer Companies	
	A. Primary Manufacturers	1,500
	B. Secondary Firms and New Start-Ups	1,000
		$ 2,500
3.	Cooperative R&D	
	A. MCC - Consortium	1,000
	B. Semiconductor Research Corporation	100
	C. Stanford University Center for Integrated Systems	300
	D. Massachusetts Microelectronic Center	100
	E. California Microelectronics Innovation & Computer Science Program	100
	F. Minnesota Microelectronics & Information Science Center	100
	G. North Carolina - Microelectronic Center	100
	H. Florida State Supercomputer Computations Research Institute	100
	I. NSF University Supercomputers Centers - Matches	300
	J. Washington-VLSI Technology Consortium	15
	K. Indiana Computer Integrated Design, Manufacturing & Automation Center	20
	L. Other	50
		$ 2,285
	TOTAL	$ 8,135

R&D for the third generation of supercomputers is different from the second generation. The second generation was in many respects an outgrowth of the things which CDC and Cray learned from their involvements in the government/university driven market for supercomputers. They invested their own funds and took all of the risk.

The third generation of supercomputers is still predominantly a private sector development. Rather than being initiated by the larger U.S. computer companies as in the first generation, they are being developed by smaller computer companies. Cray, ETA Systems, Inc. and Denelcor are by computer company standards small companies. Cray is investing 20 percent of its sales revenue in R&D. ETA is still not independently financed. Denelcor is experiencing financial problems. Each of the supercomputer companies is involved in investing $90 million of 1984 dollars or more to develop its third generation of supercomputers. Such R&D investments will be a continuing need for each succeeding generation.

All three companies are relying on basic component research to be conducted through semiconductor company in-house research or on government sponsored research programs. While most of such R&D funds will be from the Federal agencies, primarily DOD, a number of states have begun to sponsor microelectronic research generally under cooperative research arrangements between businesses and universities. The Semiconductor Industry Association has also sponsored and supported centers of excellence in component research as well as selected research projects at a number of universities.

R&D for more productive and efficient means to design and develop the components to design the architecture for the supercomputer, to manufacture and test the supercomputer, and to develop operational and application software is also being conducted across a large number of independent federal and state government agencies, manufacturers and various cooperative research programs and projects.

The largest cooperative R&D program is that sponsored by the 21 companies that comprise MCC at Austin, Texas. None of the three supercomputer manufacturers are among the sponsors to whom research and development results are released prior to general licensing three years after their availability to members of the consortium.

The establishment of the NSF University Supercomputer Centers has involved two of the three supercomputer companies; namely, Cray and ETA as well as IBM. The recent NSF created University Supercomputer Centers are utilizing the following machines --

University Center	Manufacturer's Model
1. University of California at San Diego	Cray XMP
2. University of Illinois - a new center for supercomputer research and development	Cray XMP
3. Cornell University - Center for Theory and Simulation in Science & Engineering	IBM 3084QX and FPS 164 & 264 scientific processor
4. Princeton University - John Von Neumann Center	CDC - Cyber 205 to be subsequently upgraded to ETA-10

These University centers and others at Colorado State, Purdue, Florida State and Georgia provide an R&D window for supercomputer manufacturers on current and future applications and their impact on supercomputer designs that lead to subsequent generations.

The supercomputer companies are directly involved with the Japanese supercomputer race in the marketplace. Both the U.S. response to the Japanese 5th generation computer challenge, or ICOT Project as it is now called, and the supercomputer scientific and engineering race from an R&D perspective have been taken up by private companies forming MCC and by the DARPA initiated Strategic Computing Survivability Program. To the best of our knowledge, the current three supercomputer companies are not directly involved with either program.

In addition, the U.K. and European Community have launched five-year research programs in the area of supercomputers. These programs are generally referred to as the Alvey and ESPRIT programs.

B. Current Structure of the Emerging Supercomputer Industry

As an infant industry, the supercomputer industry is in an evolutionary stage. Since my last testimony before this committee on November 16, 1983, the industry has taken on a different nature. The industry has become more diversified by broadening its base with slower machines for computational purposes. It can be described as follows:

1. Supercomputer manufacturers;
2. Superminicomputers manufacturers;
3. Add-on to mainframes; and
4. Software, graphics, and other peripherals.

Supercomputer manufacturers. Of the supercomputer manufacturers, there are three smaller companies, and one Fortune 500 company in the U.S. They are:

Manufacturers	As of June 1985 Estimated Number of Systems Delivered
Cray Research, Inc.	61
ETA Systems, Inc.	39*
Denelcor	6
Amdahl	0
	106

* Includes Cyber 205 only which is manufactured by CDC

Both Cray and ETA are located in Minneapolis. Denelcor is located in Aurora, Colorado. Amdahl is located in Sunnyvale, California.

There are also three large diversified Japanese companies in the supercomputer field. They are:

Manufacturers	As of June 1985 Estimated Number of Machines Delivered
Fujitsu	2
Hitachi	1
NEC	0
	3

There is no question that in the U.S. non-diversified firms dominate the supercomputer field. The opposite is the case in Japan. So far only Cray has reported profits. Its profits for 1984 were recorded as $45 million. Denelcor's 1983 losses were $10.3 million on sales of $3.6 million.

Superminicomputers and minisupercomuters. There are a number of manufacturers who develop and sell computers that are not the fastest for scientific and engineering purposes that require the 1 megaflop range. These computers generally sell for under $500,000. There are a small number of companies in this segment of the market. Amongst them are (1) IBM, (2) Digital Equipment Corporation (DEC), (3) Scientific Computer Systems, Inc., (4) Convex Computing Corporation, and (5) Floating Point Systems.

Both IBM and DEC are traditional computer companies. IBM has in the past responded with models to service a segment of the scientific and engineering computation market. They have made a conscious management decision not to compete in the relatively small market for supercomputers. However, they have always had at least one of their core mainframe generations available for the computing market, e.g., 370/195 and 3083. These machines sell in the multimillion dollar markets.

DEC on the other hand aggressively has pursued the scientific and engineering calculation market even though it has not entered the supercomputer markets. They have been selling superminicomputers which are typically 1/100 as fast as the Cray-2. Their minisupercomputers are the VAX 11/780 class which has an operational speed of 1.1 million instructions per second. They sell for between $125,000 and $500,000 each. The VAX 11/780 class as a superminicomputer and VAX-1 supermicrocomputer (one with .36 million instructions per second and sells for $11,245 per machine) together have accounted for over 35,000 machines sold.

In the last few years a newer type of scientific/engineering computer has appeared. This can be classified as the minisupercomputer. These machines are the result of evaluating three approaches to make slower computers to solve supercomputer applications at a lower cost. The two approaches that were discarded are namely (1) have superminicomputer run faster; and (2) take an IBM mainframe and provide it

with an add on feature for vector instructions. The third approach was selected; namely, design a totally new architecture that is faster than the superminicomputer and less expensive with a new software or software compatible with Cray or ETA. One of the current minisupercomputer manufacturers is Scientific Computer Systems, Inc. (SCS) of Wilsonville, Oregon. They made their architecture compatible with Cray X-MP after securing Cray Research permission. These minicomputers can become satellites for existing Cray users as well as reduce the Cray XMP loads by offloading to the SCS-40 which is scheduled to be available in 1986. The SCS-40 operational speed is specified to be 1/4 of the speed of the Cray-1 or 20-50 million instructions per second. The Cray XMP software is compatible with the SCS-40. The SCS-40 could be about 20-50 times the speed of VAX11/780 for approximately the same price of the high ended VAX, about $500,000.

A second minisupercomputer company is Convex Computing Corporation of Richardson, Texas. Their machine, the C-1, is a stand-alone machine that runs at about 1/4 the speed of Cray. However, they have directed their design to compete with the superminicomputers and have exploited the development of AT&T's Unix operating system to compete with DEC's VAX. They have taken a broader market segment than just the current scientific/engineering computing market of the supercomputer manufacturers. They also hope to sell their machine in the still evolving computer integrated manufacturing market. The Convex C-1 sells for approximately $495,000.

Both Scientific Computer Systems and Convex Computing Corporation believe that their next generation could well be in the same speed range as the Cray-1, or 80-200 million instructions per second, and carry a sales price of approximately $1 million. Both Cray and ETA managements have decided not to compete in the low end of the market at this time. One can safely predict that other smaller companies and startups will join Scientific Computer and Convex for a share of the lower end market.

Add-on Systems and Flexible Building Blocks. Another dimension of the supercomputer industry is not to utilize either a supercomputer or a minisupercomputer or superminicomputer. This approach is to add computational processors to existing computers. Their design, therefore, is a fine balance between the types of applications and kinds of mathematics involved. Some companies have utilized a strategy which allows them to design add-on systems processors to be attached to current mainframes or superminicomputers to handle two or three dimensional arrays of numbers as well as vectors. One of these companies is Floating Point Systems in Beaverton, Oregon. It has targeted both the VAX superminicomputers and the IBM mainframe to attach its FPS-264. When attached, the computational speed becomes as powerful as the Cray-1 or about 80-200 million instructions per second. The FPS-264 sells for $640,000. Another add-on system company is Teradata Corporation of Inglewood, California.

There are a number of companies which are developing supercomputers by linking flexible building blocks or by using a multipro-

cessor approach. Two companies in this field are BBN Labs, a subsidiary of Bolt Beraex & Newman, Inc. of Cambridge, Massachusetts and Hydra Computer Systems, a division of Encore Computer Corp. BBN has already installed 16 systems with a total of 227 processors at a cost of about $8,000 per processor.

Some experts did not consider multiprocessors as part of the supercomputer industry until more than 1000 processors were involved. Thinking Machines Corporation of Cambridge, Massachusetts has currently up and running a "connection machine" with 16,000 processors. They expect to deliver a first 64,000 processor machine to DARPA in the fall of 1985. The 16,000 processor model has established a benchmark of 250 million instructions per second or 1/4 faster than the Cray-1.

Another company in this class is Sequent Computer, Inc. with its system, Balance 800.

There are a number of R&D flexible machines being developed at more than 50 universities and start-up companies. Flexible has developed a general purpose computer called Flex/32 that performs parallel multicomputing while maintaining an "essentially unlimited expansion capability." Flex has made 4 installations as of May 1985. One was installed at Trinity Technologies Corp. in Dallas, Texas for process-control application. Another has been delivered to Structured Software Systems, Inc. of Irvine, California for satellite tracking systems. A third went to Purdue University's Center for Parallel and Vector Computation for research purposes. The fourth was delivered to

Langley/NASA Research Computation Group for numerical solutions to do with aerodynamics and structural mechanics.

Table 2 summarizes selected deliverable computer types from microcomputers to supercomputers.

In summary, the supercomputer industry in many respects is becoming more than just the firms that produce the biggest and fastest parallel processing machines. The fastest supercomputers are much like the larger mainframes for data processing. It is now evident that it is possible to build slower minisupercomputer machines that are both compatible as well as non-compatible with the faster supercomputer in terms of operations softwares. The ability and the needs to network supercomputers as well as minisupercomputers is also evident. What is not clear is what the network protocols should be. Currently network protocols in terms of networking supercomputers is at its early beginnings. Networking for supercomputers has been undertaken with success by DARPA, DOE and NASA. The next round of evaluation and implementation of networks will be funded by NSF. How these networks will be compatible to handle the several generations of Crays, CYBER 205's. ETA, Denelcor and Amdahl supercomputers is still too early to determine.

The other networking aspects are the ability to network minisupercomputers with the supercomputers. The needs are not difficult to state. They will be required to enhance individual scientific research as well as comparative research between scientists in dif-

ferent geographical locations. They will be needed to increase the effectiveness and efficiency of educating scientists, engineers and other graduate professionals in business, economics, and other social scientists. The minisupercomputer when shared with the main supercomputer can help to encourage both research activities as well as the education and training of the next generation students. The various networking needs such as graphics and data bases are not as far along for supercomputers for scientific research or education.

Another aspect of networking is that connected with computer integrated manufacturing (CIM) concepts including VLSI manufacturing. While the supercomputer industry can play an important role in CIM, it is still in its early stages. The barriers are standarization of network, cost, lack of expertise at top management levels and a shortage of systems engineers and other trained personnel to implement CIM. However, the fact that we are seeing the supercomputer industry emerging and evolving into minisupercomputers, flexible parallel computers and add-on systems is encouraging. This indicates that there are more dimensions to the marketplace than governmental scientific/engineering computations, university scientific/engineering research and training, and selected industrial research and design.

C. Applications, Markets and Trends

The applications for supercomputers are still those that I presented in my testimony of November 15, 1983. The following quotation from that testimony is appropriate:

TABLE 2

Selected Micro to Supercomputers
(speed and price)

Computer Type	Company	System Name	Speed in MIPS (*) Megaflops (+)		Price in $
Microcomputer	Apple	Apple	.0005	*	$ 1,795
Supermicro-computer	Digital Equip. Corporation	Micro VAX I	.36	*	11,245
Minicomputers	Digital Equip. Corporation	PDP-11/44	.4	*	29,950
Supermini-computers	Digital Equip. Corporation	VAX-11/780	1.1	*	125,000 to 500,000
Mainframe	IBM	IBM 3083	4.2	*	735,000 to 1,975,000
Minisuper-computers	Scientific Computer Systems, Inc., Wilsonville, Oregon	SCS-40	20-50	*	about 500,000
	Convex Computing Corporation, Richardson, TX	Convex C-1	60	*	495,000
Add on Systems	Floating Point Systems, Beaverton, OR	FPS-264	27-67	*	640,000
Supercomputers	Cray Scientific Research	Cray 1	80-200	+	7,200,000
		Cray 2	800-1400	+	17,000,000
		Cray 3	10,000	+	?
	ETA Systems	Cyber 205	800	+	6,000,000 to 15,000,000
		ETA 10	10,000	+	9,000,000 to 20,000,000
	Amdahl	Vector 1100	267	+	7,700,000
		Vector 1200	533	+	10,700,000
Flexible Parallel Processors	Thinking Machines Corporation	connection machine	250-1000	*	?

The supercomputer is the single greatest impact on world communication, automated factories, health care delivery, biotechnology production, renewal of basic industry and heightened productivity of the service industry, including government. The real task is how to develop appropriately integrated policies, regulations and support mechanisms that extend the U.S. computer/communications industry. The commercialization of supercomputers for the global market is so tightly structured from scientific exploration to ultimate use, regeneration time is so short, investments so large and risk so great, that we cannot leave policy consideration to evolve accidently and independently as in the past. The computer/communications industry is our future source of large scale employment; i.e., one out of six. Table 3 emphasizes the importance of the supercomputer as a driver for the computer/communications industry.

The consequences of losing economic and scientific preeminence in the supercomputer industry are vast. The supercomputer is a central driver for the rapidly emerging worldwide computer/communications industry. It impacts communication developments, the renewal of basic industries, productivity increases and the development and expansion of new industries. It is essential in improving our educational structure, fulfilling critical manpower requirements and enhancing our industrial creativity and innovation. It is the seed for encouraging the emergence of a myriad of technology venture businesses in the context of a private enterprise system that has always been the unique American way to achieve and maintain U.S. economic and scientific preeminence.[5]

Applications for the fourth and beyond generation supercomputers are primarily for research and education. The research markets are the government labs and industrial research laboratories and academic research. The research and education market has gained momentum particularly during the last year and one half. The stimulus in many respects came from accomplishments in the government and industrial laboratories applications for aircraft aerodynamics, three-dimensional modeling of oil reserves, emergency shutdown of nuclear reactors, utility power grid planning, molecular analysis, structural analysis for automobiles, shipbuilding and skyscrapers, computer-generated imagery

TABLE 3

Effects of Preeminence in Supercomputer Technology
on the Computer/Communications Industry

Drivers	Area Affected

Industries Developed and Expanded by Super Computer Related Technologies	Robotics/Industrial Automation
	Computer-aided Engineering
	  Computer-aided Design
	  Computer-aided Manufacturing
	  Computer-aided Testing
	  Computer-aided Quality Control
	Biotechnology
	Office Automation
	International Financial Services
	Service Industry
	Computer Industry
	Semiconductor Industry

Revitalization of Basic Industries	Energy
	  Petroleum
	  Nuclear
	Steel
	Automobile
	Textile
	Chemicals

Institutions Receiving Significant Productivity Boosts	Government Defense Education Research
Communications Technology Enhancements	Subscribers Equipment Telephone Co. Equipment Overseas Carriers Satellite Carriers

for movies, medical diagnosis and product design, computer-aided design manufacturing and testing and others. These successful applications are currently moving into business and industrial applications.

There are a number of other research and educational applications that require the capabilities of third generation and succeeding generations of supercomputers. These applications are what Neil Lincoln of ETA Systems, Inc. defines as leading-edge efforts in science and engineering which cannot be solved by the currently available supercomputers. More simply put, there are applications that are at least a generation ahead of the third generation supercomputer capacity/abilities. These applications include the needs for the physical scientist, engineering of large scale programs and processes, management of large scale projects and programs, long term economic modeling at global, national and local community levels and others. It is perhaps important to stress that the research and educational market has a critical need for supercomputers that are of the n^{th} generations -- beyond those in development. These are machines required for research in knowledge procesing or what the Japanese called the 5th Generation. Knowledge processing machines will provide the momentum for future research efforts to be transferred and diffused for applications in the service, business and industrial markets. Ultimately, there is a need for intelligence processing which is beyond knowledge processing in that these machines will make it possible for new knowledge to be acquired from utilization of

53-831 O -86 - 9

large-scale knowledge bases (yet to be organized, built and maintained) through computer learning.

The applications in business and industry are for the aerospace, automobile, energy, and chemical industries. The firms in these industries can be characterized as large and capital intensive. They have been using supercomputers for developing specific applications such as wing design, simulation of in-flight conditions, seismic data analysis, designing new synthetic materials, genetic engineering, and total factory automation evaluations. These industries are becoming the market for second and especially third generation supercomputers as well as for minisupercomputers, flexible parallel processors and add-on systems.

The market for supercomputers particularly in the last eighteen months has been a surprise to manufacturers and others. The surprise is that it not only has exceeded the heretofore early projection for a maximum world market of 100 supercomputers but that Cray alone has sold over 100 actual processing machines. The total number of sales by systems - (some may contain more than one central processor) for U.S. manufacturers exceeded 106 by June 1, 1985. All four of the U.S. manufacturers have backlogs for deliveries into 1986-87.

Of equal importance is that the market mix has shifted from government labs to more industrial and university sales. Currently, it is estimated that the government lab markets are about 35 percent of the total market.

There is a need to segment the markets and applications for supercomputers for policy purposes. One segment is the research and education market. Another is the business and industrial applications market. Each segment's needs and applications are totally different. Yet they are interlinked and supportive of each other, particularly in developing the overall supercomputer industry including minisupercomputers, add-on systems, flexible parallel computers, graphics, software, etc.

The research and educational applications and their markets are also expanding. Earlier applications and market demands were largely federal government driven. The current shifting of NASA efforts from the shuttle to the space station, of DOD to strategic defense initiatives, of DOE from alternative energy research to nuclear defense and weapons are all major programs that will require supercomputers. The national physics supercollider program will also require a large investment in supercomputers.

One should applaud the Supercomputer University Center Grants by the National Science Foundation (NSF). These grants are important steps in helping to ensure the overall scientific preeminence of the U.S. particularly in the area of particle physics, biotechnology, aerodynamics, weather monitoring and prediction, etc. These grants together with other U.S. university supercomputer centers at Georgia, Florida State, Purdue and Colorado State and others who are Non-NSF and DOE dependent no doubt will provide an important thrust towards

ensuring U.S. preeminence while pushing the frontiers of U.S. basic research. Their research will help to expand the supercomputer industry particularly for minisupercomputers, graphics, artificial intelligence, CAD/CAM, flexible manufacturing, robotics, knowledge base systems and others. These grants, as expressions of a national policy, can begin to build important institutional links and informational networks systems between:

- o Academia and industry.
- o Scientific laboratories and centers doing complementary work.
- o Scientific centers which require interactive scientific-technological dissemination.
- o Researchers in academic and industry.
- o Federal government, universities and corporations.
- o Federal government, state government, universities and corporations.

These linkages are relatively new. We do not have much experience with them; yet, we cannot let important policy considerations evolve accidentally. The University Supercomputer Centers are primarily established to ensure U.S. scientific preeminence as well educate scientists and engineers for supercomputer usage. They have yet to be linked for maximum market and application purposes.

These current university-based centers are generally involved with basic research. The next 5-10 years will see the traditional conduct of scientific and engineering research transformed to more computer

simulation with less experimentation than required in the past. These shifts are taking place for several reasons. First, the critical needs of our nation such as defense, energy and health are increasingly more multi-disciplinary. Second, the research is becoming more complex and costly. Third, the science disciplines are increasingly interrelated and overlapping when addressing critical needs for new technologies. Fourth, there is increasing recognition that newer technology developments that are funded by and with government sources have an implied commitment to their short-term transfer for the purpose funded as well as longer term diffusion for other economic benefits and positive impacts. The transfer and diffusion of government supported R&D, especially for supercomputers and related research, has not been available since the first generation. As stated earlier, the second generation was primarily funded by the private sector.

It is important to understand how transforming the way basic research is conducted can impact current policies. There is a need to reexamine the traditional approach to intellectual properties and their dissemination. The issues include more than ownership, which is complex enough given the expanding network of relationships and competing claims. They include ensuring the vitality of and need for the centers and developing adequate provisions for transfer and diffusion of their results.

In the past, little attention was paid to how science was transformed into technology which was subsequently transferred for specific

commercialization purposes and then diffused throughout all industries, regionally as well as for international trade. The general paradigm was that basic research innovations would be utilized for applied research and developments and that their manufacture would naturally follow. Diffusion to other uses and industries would occur when R&D results were both economical and better understood in general. The utilization of technology as a resource was perceived as an individual institution's responsibility. Economic developments flowed from this process because of individual firms' ingenuity and their entrepreneurial spirit. Targeting may have been a Japanese national policy, but in the U.S., market opportunities at home and abroad seemed sufficient for economic growth and diversification. It was expected that all regions of the U.S. would in time enjoy the benefits of this paradigm in which new innovations from research were followed naturally by timely developments, commercialization, and diffusion.

The current experimental and collaborative uses of supercomputer centers are a break in the traditional way of doing research and perhaps subsequent commercialization. The major breaks are in the following areas:

1. Financing is a collaborative effort. The other financial supportes are (1) companies that develop computers, (2) users of supercomputers and research results, (3) state governments and (4) individual universities that comprise the consortium.

2. The research results are expected to be more basic research than proprietary products for industry. Both software and newer supercomputer designs are expected to emerge from the various centers and their participants.

There does not seem to be a built-in mechanism for self sustaining the centers. If they are predominately doing basic research or advanced research for complex, multi-disciplinary research, their end products are more publications of research results and less tangible goods and services. Publications and electronic transfer of research results may be important educational frontiers, but they may not be sufficient output for the centers' continued self support. The replacement of currently selected supercomputers will continue to be capital intensive. Maintaining and expanding software will be expensive, e.g., at least one-third of the cost of the supercomputer itself. Long-term funding will be a problem. What sources of research support for the centers will become available? Will they continue to depend on cooperative support?

The cooperatively funded supercomputer centers are primarily newer institutional developments of federal government, state government, business and universities. Today the new institutional developments emerging around the supercomputer centers focus on maintaining U.S. preeminence in scientific research. Their role and scope as a policy matter at present is limited to five years of federal support. There are no built-in mechanisms to support these centers in the long run nor

their longer term roles for research, economic growth, employment generation and economic diversification. Assuring that these centers do not become anachronistic institutions to be saved is a major policy matter. This changes the role of the federal government especially in cooperative R&D and its subsequent commercialization for markets and applications for both the supercomputer industry and those upon which they impact.

In my opinion the NSF sponsored University Supercomputer Centers, as well as federal government laboratories are the leaders in trying to implement an emerging policy; that is that research and developments that are funded by and with government and other institutional funds have an explicit commitment to technology transfer and diffusion. Federal government laboratories are in the forefront of supercomputer research and application. They are now seeking ways to transfer technology and ensure its subsequent diffusion. The IC^2 Institute at The University of Texas at Austin has been in close touch with Los Alamos National Laboratory to assist in the process of implementing effective technology transfer mechanisms. Los Alamos Lab has provided a number of workable newer institutional arrangements. Among these are arrangements for meeting places for people with common research interests such as the Center for Material Science, Center for Non-Linear Studies, and a branch of University of California's Institute for Geophysics and Planetary Physics. This also includes workshops for those in interested industrial firms and arranging for joint projects with other nations' participation. Dr. Donald M. Kerr,

Director of Los Alamos National Laboratory has summarized the needs for effective transfer and diffusion as follows:

> Finally, we try to recognize that both our very applied programs as well as our technology development programs have a need for people who are not purely disciplinary in nature, but in fact can play the new role of integrator, mediator, catalyst, or translator, whose technical breadth is required to work on these multidisciplinary activities, whose depth of knowledge has to be sufficient for them to be effective. We recognize that such people are more often found in the work place and trained on the job than produced through a traditional scientific or engineering education.
>
> We see a long-term need to find a mechanism to encourage these people; to identify them, to give them the opportunities to develop their capabilities as they move from small to larger programs, both in relatively basic research as well as in the more applied development activities. This we think, coupled with the way in which science is increasingly done, taking advantage of large-scale computer simulation, will provide for effective and efficient research and development in the future on a timescale commensurate with national needs. The National Laboratories can adjust to meet the challenge I stated at the beginning, but it requires more effort than in the past, more willingness to experiment with new structures and processes, and more willingness to reach across boundaries to create partnerships with industry, financial institutions, and universities. The goal is to continually renew a creative anachronism to focus on multidisciplinary problem solving, enhance productivity, and create a pool of talent for the future.[6]

Transfer and diffusion require more than government laboratories or the university center's efforts. There must be institutional developments at state and local government levels and in the private infrastructure that assure such transfer and diffusion in meeting market needs and demands in a timely manner with fair distribution to all regions of the U.S. There is a need for institutions to provide seed capital for economic utilization of the centers' research outputs

for commercialization which cannot be the responsibility of the government laboratories and universities. Innovation in the United States is primarily a small-and medium-size company phenomena. Currently, the exploitation of supercomputers is primarily by the larger companies. This is true at the present for the firms in the aircraft, automobile, and chemical industries. The stimulation of innovation through supercomputers for small companies today is a decided gap in policy formulation.

Current market trends for annual worldwide supercomputer sales indicate that sales on a units-sold basis are on track according to IC^2 Institute projections in November 1983 for the Committee on Science and Technology, U.S. House of Representatives. Chart B shows the early projections. The projections for 1984 for annual supercomputer sales in billions of dollars are higher than actual sales recorded by the industry by about 25 percent or $100 million. The projections did not take into account lease sales as such but projected them as units sold. The November 1983 projections scenarios are shown in Chart C. The projections for supercomputer sales shown in Chart C are for the conservative scenario. This scenario has been adjusted to account for sales from minisupercomputers, add-on, other peripherals and software.

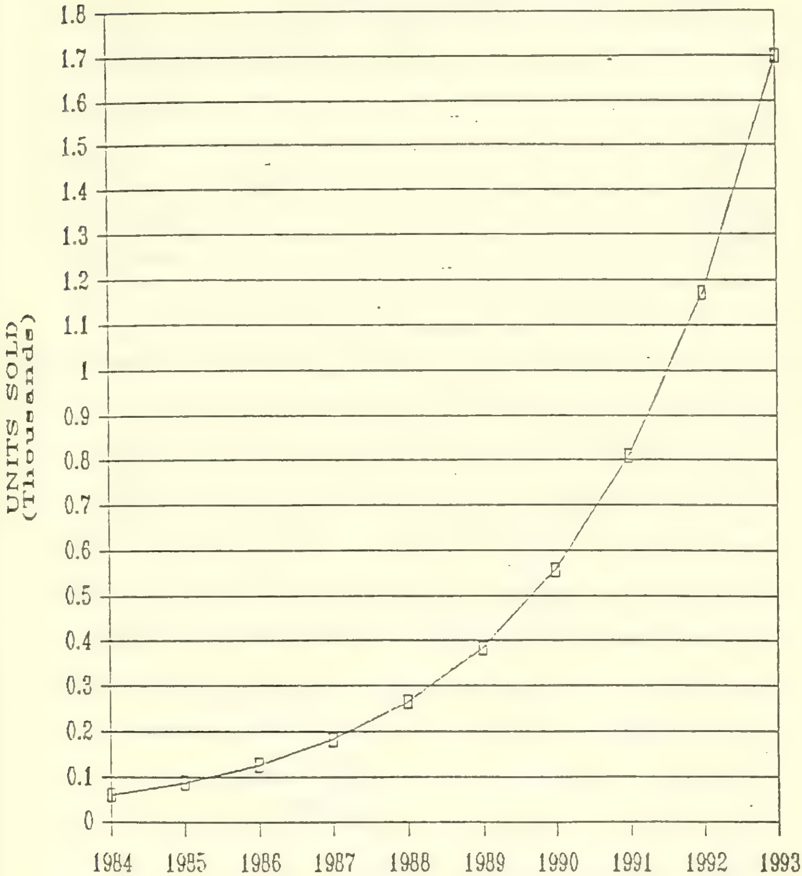
Trends in the supercomputer industry can be summarized as follows:

1. There is an increasing demand for supercomputers that was unforeseen as short as eighteen months ago. Demands are

Chart B

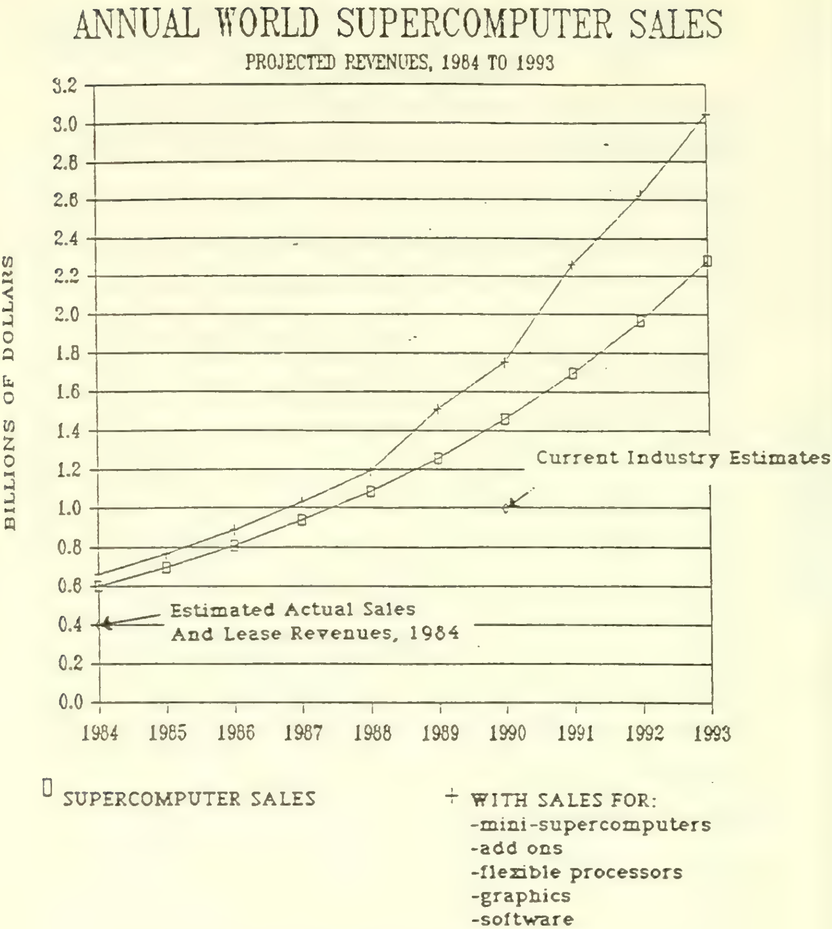
ANNUAL WORLD SUPERCOMPUTER SALES

PROJECTIONS FOR 1984 TO 1993



Source: IC[2] Institute, The University of Texas at Austin.

Chart C



Source: IC² Institute, The University of Texas at Austin.

arising for federal government needs, expanding industrial and business markets (domestic and foreign), and universities (domestic and foreign).

2. All four U.S. supercomputer manufacturers are reporting a relatively robust firm backlog.

3. The smaller supercomputer manufacturers are finding a developing market for their products in education, research labs as well as business and industry. There are strong indications that their next generation in development will broaden the number of applications and their markets.

4. Software firms for supercomputers are expanding and beginning to deal with emerging applications.

5. University Supercomputer Centers are focusing on advancing research and education in supercomputers.

6. University Supercomputer Centers, particularly consortiums at Princeton and the University of California at San Diego, can advance the communication networks between the collaborative supercomputer users. These advanced networks can be 3-8 years ahead of the networking of large scale, data processing computers.

7. The supercomputer network can and should avoid many of the problems that have beset the local area and long distance network problems for data processing and office automation.

8. There will be an increasing number of collaborative ties with the aim being to transfer and diffuse technology. These ties will include state governments, federal government, large and small companies, and universities.

9. International competition will become more fierce, especially from the Japanese in the next year. It is still too early to predict the competitive impacts from the ESPRIT program of the European Common Market and the United Kingdom's Alvey Program.

Part II. Discussion of the Issues

1. How will supercomputer R&D continue to be performed and funded in the future? What are the changing patterns?

For initial supercomputers and the first generation of supercomputers, R&D was funded by the industry and government. The first customers for the deliverable second generation supercomputers were government labs. For the third generation of U.S. supercomputers, R&D is being funded by industry. This is the generation about which the Japanese announced that they would be ready to enter into world-wide competition by 1990.[7] This program of R&D is funded at $400-500 million jointly by the Japanese government and six major, diversified firms in their computer industry.

The first known customers for U.S. deliverable third generation machines are government labs and university supercomputer centers

funded jointly by the federal government, state government and computer manufacturers.

For the fourth generation of supercomputers, R&D is funded by the federal government. They are being developed by DARPA and the Livermore National Laboratories. The fourth generation U.S. specifications are substantially ahead of the current Japanese 1990 goals and are in many respects a way of meeting the Japanese challenge. To the best of our knowledge, none of the four supercomputer manufacturers for the second and third generation supercomputers is actively involved with these R&D efforts.

Both funding and performing R&D for future supercomputers will be conducted under a number of different patterns. Among these are:

a. Continue R&D with non-government funds.

b. Enter into joint ventures with other companies, including foreign firms and licensing newer technologies.

c. Participate in cooperative research, non-profit consortium such as MCC for advanced research for enabling technologies to be selected for implementation by each participant.

d. Enter into R&D contracts with federal agencies for a specific supercomputer or enabling technologies that can be used later for developing commercial supercomputers.

e. Sell equity to other corporations that also license for their purposes the technologies being developed for supercomputer hardware and software.

f. Utilize R&D partnerships and other traditional capital venture processes.

g. Have government labs and/or universities buy or lease early models of next generation supercomputers for help to finance the supercomputer developers as well as use them as beta sites and thus securing through operational time for bank and/or leasing financing.

h. Establish collaborative links with research being performed by governmental laboratories and universities. These links can be to utilize relevant intellectual property funded by others such as the federal government or state governments or directly funded by gifts or business donations. Other collaborative linkages would be to sponsor university research for commercialization development and then retain all of the commercial rights.

2. What kinds of complex applications and needs will drive the advanced developments of supercomputers?

In the past a number of major computer manufacturers had dropped out of the supercomputer industry for a number of reasons. The primary one was that the market was not established nor as profitable as that for data processing and automated office. The challenge from the Japanese was a major impetus for development of the U.S. fourth generation supercomputers. Yet the second and third generation manu-

facturers have met the Japanese challenge. The emerging and still critical marketplace for the supercomputer world is being established by advanced government applications that are transferable into the supercomputer business and industrial market segments.

The current transfer and diffusion mechanisms need to be assessed to determine how to develop a market when there is not a replacement or when the substitution marketplace is required. Building a market using newer technology generally requires an infrastructure of specific needs that may not exist or be anticipated. In these respects, we know that there are a series of complex problems and special needs that even the fourth generation supercomputers cannot solve. Among these are:

a. Large scientific research projects. Livermore Laboratories have already identified problems in nuclear weapons research that take 500 - 1000 hours to solve on second generation supercomputers. There are other such complex computational problems and research areas that will lead to Nobel prizes as well as newer markets and perhaps industries. In addition, knowledge processing and knowledge systems for large scale programs for scientific research, for space defense and commercialization, for scaling up newer manufacturing processes and for biotechnology are beyond the third generation supercomputer capabilities.

b. Advanced flexible manufacturing and processing computerized systems. These are beyond the third generation's capability com-

putationally, as well as in terms of communication and networking among computers. There is no currently acceptable means of establishing standardization goals for more effective and efficient developments in a non-dominant competitive environment. This is the near term expandable market for the supercomputer industry, and it includes communications, robotics, CAD/CAM, CIM, graphics, expert systems, etc.

The required data bases for these applications will take a long time to develop, i.e., 5-12 years. There is no policy other than leaving it up to each user to use their discretion even for standards data. Many of the applications, including biotechnology, medical and health care and financial services, are daa base dependent.

c. Advanced security needs and applications these are still continuing as a supercomputer market. The needs for SDI, space and overall national security programs provide a substantial market for both hardware and software developments. Government needs also include supercomputers for weather forecasting, advanced air traffic control and advanced data processing.

The critical question is to determine if it is possible to develop priorities so that succeeding generations of supercomputers will not face a boom or bust cycle or reactive crises/responses. It is essential that we truly become anticipatory if we are to maintain U.S. economic and scientific preeminence.

3. Can cooperative research reduce the development time and maintain the U.S. competitive position in the global supercomputer marketplace?

Recent institutional developments for cooperative research and development including supercomputer applicability have been initiated primarily for economic growth, job creation and diversification. These collaborative efforts include industrial consortiums, software consortiums, university consortiums, and state/industry cooperative programs with universities. Government labs have already been directed and are actively pursuing collaborative efforts to transfer and diffuse their technological advances. Yet it is highly unstructured. How policies are formulated will go a long way towards making and securing our nation's and individual futures as well as the future of the supercomputer industry.

I'd now like to digress for a short historical overview to focus on this issue. Prior to 1979, there was little evidence of technology venturing, that is, collaborative or institutional developments for economic growth and diversification. The prevalent attitude was a "go it alone" philosophy that was reflected in a variety of ways. The emphasis was on industrial relocation rather than on building indigenous companies; separation of institutional relationships, especially between universities and corporations; adversarial roles between government and business; and reactive rather than proactive policies both nationally and industrially to met international competition.

We, at times, seemed to believe that the rules of the game were set in concrete rather than subject to the dynamics of an ever-changing, global environment and to an economy that was coupled to changing values. As well as the shifting roles of the university in research and education.

A short six years ago, technology and its impacts were more threats than opportunities with which to build a future. Six years ago, total annual venture capital was less than the then current one-day's loss of Amtrak operations. Entrepreneurship was ignored as a force or driver. Technology transfer and diffusion were subjects for research and not a mandate for commercialization of research and development. Six years ago, there was little doubt about U.S. leadership in high technology, particularly in terms of electronics and its industrial and scientific markets. There was no question but that "hi-tech" was a major contributor to the nation's trade balances. Six years ago, the loss of earnings, layoffs and production curtailments were not part of management's major concerns in high tech firms.

For much of the period from the 1950s to the 1980s, it was generally assumed that scientific research would in one way or another transfer into developments or technologies and subsequently be commercialized. For much of this period, little attention was paid to how science was transformed into technology which was subsequently transferred for specific commercialization purposes and then diffused

throughout all industries, regionally as well as for international trade. The general paradigm was that basic research innovations would be utilized for applied research and developments and that their manufacture would naturally follow. Diffusion to other uses and industries would occur when R&D results were both economical and better understood in general. The utilization of technology as a resource was perceived as an individual institution's responsibility. Economic developments flowed from this process because of American ingenuity and our entrepreneurial spirit. Targeting may have been a Japanese national policy in this period; but for the U.S., market opportunities at home and abroad seemed sufficient for economic growth and diversification. It was expected that all regions of the U.S. would in time enjoy the benefits of this paradigm in which new innovations from research were followed naturally by timely developments, commercialization, and diffusion.

The current 8 University Supercomputer Centers and those yet to come are a good entry point for this issue. Government laboratories have begun to transfer and diffuse technology. They have identified a series of needs and especially persons who can perform the role of integrator, mediator, catalyst or translator and whose depth of knowledge is both technically, financially and managerially adequate for them to be effective. These persons are not produced through traditional scientific or engineering or managerial education. Universities have just begun to formulate policies and establish more effective organizational structures to handle intellectual properties

more economically. This is not enough for what we are talking about under this issue. Professor William B. Rouse has stated this issue as follows:

> "Despite the fact that basic research occassionally produces better mousetraps, the applied world seldom visits the laboratory door...The issues are many and complex: they are more organizational in nature than technical."
>
> "Specifically, the university evaluation system should reward involvement with real world problems. Further, interdisciplinary cooperation should be vigorously encouraged; current academic evaluation and reward systems, at best, only tolerate such cooperation. Perhaps the best place for such changes to begin are in professional schools such as medicine, law, and engineering where real-world involvement would seem to be natural. Unfortunately, for these changes to occur in engineering, for example, the tendency to emulate physics and mathematics will have to be reconsidered.
>
> The university's relative immunity from "market forces" make such changes feasible. For the same reason, organizational changes within the university are very slow. Thus, a short-term strategy for fostering technology transfer is also needed. It seems to me that any group of reasonable persons should be able to agree that new and/or vastly improved mechanisms are needed for transforming information in the research base into an appropriate form for the applications base.[8]

The Japanese seldom use academia in the actual targeted projects. They more often utilize business and government laboratory personnel. They also establish time frames for the project much like our"sunshine laws." They also accomplish transfer and diffusion when the researchers return to their firms or laboratories.

This is not necessarily what is required in the U.S. Our social and work values are different. We do not utilize lifetime work principles of employment in either our firms or government laboratories.

The newer supercomputer centers can modify and promulgate newer approaches to better enable their research to be more rapidly and efficiently incorporated in applied ways by firms. With this approach the past time span of 15-30 years to transfer technology into economic products and services can be substantially shortened. The utilization of intellectual property rights to better diffuse research results is also a good possibility. However, without explicit policy to address this issue, the traditional approaches will prevail.

4. How can the results of cooperative research and supercomputer technology be transferred and diffused regionally?

With government labs and university supercomputer centers as a prime focus we can better illustrate this issue. It applies to other cooperative research, institutional developments that involve federal, state, and local governments, along with businesses and academic institutions in collaborative efforts. When national laboratories or centers are established, they produce among other things national technological resources. How to distribute these equally or fairly among the regions of our country is a newer issue. In the past and especially under full employment, this was not an important issue. Since 1980, various states have been concerned with unemployment, job creation and the utilization of high technology. Over 150 initiatives have been taken over the past five years by the states.

The IC^2 Institute at The University of Texas at Austin has conducted a study to determine where the innovative developments are

taking place. For these purposes, the major drivers used were Federal R&D obligations, traditional venture capital, and selected company R&D expenses. The dominant states by ranking are shown in Table 4.

At present, there are 17 states and the District of Columbia that meet the innovative criteria. Four states are ranked in the top 10 within each category; namely, California, New York, New Jersey and Massachussets. These certainly are the first tier states. The second tier or those in at least two of the categories' top 10 would include two states: Maryland and Texas.

There are active developments taking place in restructuring firms and industries in terms of acquisitions and mergers. The dominant states, in rank order, are California, New York, Texas, Illinois, Florida, Pennsylvania and New Jersey.

While a transformation is taking place, in terms of innovation, it has become evident that the present traditional paradigm that science and technology naturally evolve into commercialization that makes and secures a nation's future is inadequate. It does not adequately provide employment opportunities; mitigate layoffs; maintain a strong global competitive position; make economic security; and present growth opportunities across the board. Furthermore, the mechanism of allocating resources needs to focus more on flexibility and adaptability than on efficiency and effectiveness.

5. How can supercomputer advances be transferred and diffused to small and medium-sized firms?

Supercomputers as well as minisupercomputers are in many respects expensive and beyond the reach of many small and medium-sized firms. Such size firms are already noted for their innovative abilities and as a major source for employment growth. As supercomputer developments promulgate, they will be among the 1st to utilize their advances especially in design, manufacturing, quality control, etc. In some sense, if they could be placed in a position to utilize these results, they could help stem the flow of jobs outside the U.S. and at the same time increase other employment opportunities through new product development and increased productivity.

Currently there has been little concern or organized effort to examine this need. This aspect could be achieved by incorporating some measures that utilized the university supercomputer center network system. This would help in the development of standarization for supercomputer hardware, peripherals, station terminals, graphics and software. If done properly, it could help the growth of the emerging supercomputer industry.

Part III - Conclusion

The supercomputer industry is in its infancy. It is a promising period of growth. Its product lines are becoming diversified in terms of superminicomputers, software and other peripherals.

Today, it is an industry principally dominated by small, non-diversified supercomputer companies. They have managed to meet the

TABLE 4

Innovation State Rankings

State	Dominant Federal R&D Obligations[1]	Dominant Selected Company R&D Expenses[2]	Traditional Venture Capital[3]
California	1	3	1
New York	3	1	2
New Jersey	8	3	10
Massachusetts	4	5	3
Maryland	2		6
Texas	7		8
Illinois			5
Virginia	5		
Florida	9		
Nex Mexico	6		
District of Columbia	10		
Michigan		2	
Pennsylvania		5	
Connecticut			4
Colorado			7
Minnesota			9
Delaware		5	
Ohio		5	

[1]For fiscal year 1983

[2]1983-84

[3]1984

Source: IC^2 Institute, The University of Texas at Austin.

Japanese challenge. They are successfully dominating the global supercomputer market. The long term future for the supercomputer industry is still promising. It depends on how the user industries markets are developed. These markets are not necessarily replacement or mainly substitutions for current obsolete equipments, methods and services. The markets are here now and can meet the visions for the new American dreams.

This industry is also pluralistic. It provides the key tools for basic research and advanced engineering. It will change the nature of higher education and may well provide the required thrust for longer term means of financing academic research based on newer institutional arrangements under technology venturing.

Yet it is not a foregone conclusion that there are no additional policy issues that need to be attended to. There are a number of significant issues that need the attention of this committee. Chief amongst these are to assure that the university supercomputer centers are networked among the 30 university consortium members as well as between the eight centers. This technological challenge, when met, can do much to keep this nation ahead of its foreign competition in Japan, U.K. and the European economic community as well as the USSR and China. It is my prediction that we can gain three to eight years on them.

The networking, when coupled with the forthcoming U.S. third generation supercomputers, can move the supercomputer industry from its

infancy to the takeoff stage. It can also bring with it the CAD/CAM, Robotics, expert systems, and computerized flexible manufacturing industries into their adolescence and once more restore American manufacturing to world leadership.

This committee can do much in formulating appropriate science and technology policy to assure that American small business and innovative entrepreneurs are assured of the availability of supercomputer technology at an earlier date than our traditional process of commercializing research has previously made possible.

Footnotes

[1]Hwang, Kai, "Multiprocessor Supercomputers for Scientific/Engineering Applications," Computer, June 1985, pp. 57-73.

[2]Fernbach, Sidney, "Supercomputers - Past, Present, Prospects," Future Generations Computer Systems, July 1984, pp. 23-38.

[3]Thorndyke, Lloyd, Unpublished remarks to ETA Systems, Inc. Board of Directors, 1985.

[4]"Minding Everybody Else's Business," The Economist, May 4, 1985, p. 65.

[5]Supercomputers. Hearing before the Committee on Science and Technology, U.S. House of Representatives, November 15-16, 1983, pp. 92-94.

[6]Kerr, D.M., "Can Anachronistic Institutions Be Saved?" Paper delivered at IC^2 Institute of University of Texas' Conference on Frontiers In Creative and Innovative Management, Miami, Florida, November, 1984.

[7]Please note this is not the 5th Generation announcement by Japan. They have identified two target markets -- one for supercomputers for scientific/engineering computations and the other for knowledge and intelligence processing. The first is the supercomputer program, and the second is the 5th Generation Computer Program.

[8]Rouse, W.B., "On Better Mousetraps and Basic Research: Getting the Applied World to the Laboratory Door." IEEE Transactions on Systems, Man, and Cybernetics, January/February, 1985, p. 7-8.

Mr. FUQUA. Thank you very much, Dr. Kozmetsky, for a very fine thought provoking, informative presentation.

Because we are running late, we will forego the questions and we may have some that we want to submit to you in writing for you to respond to.

At this time, the committee stands adjourned.

[Whereupon, the committee was adjourned at 12:40 p.m.]

APPENDIX

Additional Material for the Record

NATIONAL SCIENCE FOUNDATION
WASHINGTON D C 20550

SUPERCOMPUTER FACT SHEET

Cornell University will establish a supercomputer facility as part of an overall Center for Theory and Simulation. The initial configuration in the Facility will be an IBM 3084QX mainframe hosting a number of FPS-164 and FPS-264 Array Processors. Cornell plans to build an early example of a 1990's computing environment to serve the advanced computational needs of graduate education and research. Costs will be shared by the University, the State of New York, IBM, and Floating Point Systems.

The University of Illinois will establish a Center for Scientific and Engineering Supercomputing. The Center will be a basic research facility serving the national research community and modeled on the successful supercomputing centers at the national laboratories. Illinois will use a CRAY X-MP/2400 that will combine multiprocessor technology with high speed vector pipeline architecture and a hierarchical memory structure. Cost sharing will be forthcoming from the University, the State of Illinois, and Cray Research, Inc.

The Consortium for Scientific Computing (CSC), a non-profit corporation of twelve research institutions, will locate its John Von Neumann Center (JVNC) near Princeton, New Jersey. A CDC Cyber 205 supercomputer will be the centerpiece of the initial hardware configuration, with a planned upgrade to the Class VII ETA 10 supercomputer when available. The JVNC will establish network communication lines and gateway hardware at each consortium member's site, and will provide remote access to all other qualified researchers. The Consortium members, Control Data Corporation, and the State of New Jersey will be sharing the costs of this facility.

The San Diego Supercomputer Center (SDSC), managed and operated by GA Technologies, Inc. (GA), will be located on the campus of the University of California at San Diego (UCSD), at La Jolla, California. The SDSC, a consortium of eighteen institutions, will be built around a CRAY X-MP/4800 machine. The system is modeled after the Lawrence Livermore National Laboratory's successful Magnetic Fusion Energy Computer Center. SDSC will also provide remote access to other qualified users. Funding will also be provided by the State of California, Cray Research, Inc., and consortium members.

DOE/ER-0218

THE ROLE OF SUPERCOMPUTERS IN ENERGY RESEARCH PROGRAMS

Published: February 1985

U.S. Department of Energy
Office of Energy Research
Washington, D. C. 20545

CONTENTS

PREFACE

On July 9-11, 1984 a meeting was held at the National Magnetic Fusion Energy Computer Center, of the Energy Research Users Advisory Committee to the NMFECC. The purpose of the meeting was to prepare this report on the role of supercomputers in ER programs and the need for more powerful computers.

The participants in the July meeting at NMFECC were

A.E. Brenner, Fermilab
J.S. Cavallini, DOE, Washington
J.M. Fitzgerald, National Magnetic Fusion Energy Computer Center
C. Grebogi, University of Maryland
W. Herrmannsfeldt, Stanford Linear Accelerator Center
P. Hough, Brookhaven National Laboratory
M. Kalos, New York University
J. Killeen, National Magnetic Fusion Energy Computer Center
W.A. Lester,Jr., University of California, Berkeley and Lawrence Berkeley Laboratory
P. Messina, Argonne National Laboratory
G.M. Stocks, Oak Ridge National Laboratory

An earlier meeting of the Committee was held on March 21-22, 1984, and in addition to the above members other participants were

S. Kowalski, Massachusetts Institute of Technology
U. Landman, Georgia Institute of Technology
S.G. Louie, University of California, Berkeley

In addition to the above participants, written contributions were received from

G. Maggiora, University of Kansas
P. Reynolds, Lawrence Berkeley Laboratory
J. Petke, University of Kansas
E. Kostelich, University of Maryland

In addition to providing the material for Chapter 4 of this report the Committee at the July meeting drafted Chapter 1 and unanimously agreed on the Recommendations contained in that chapter.

The particpants wish to thank Louise Beite and Rebecca Nelson for the preparation of the many contributions during and after the meetings.

John Killeen
Director
National Magnetic Fusion Energy
Computer Center

November 1984

1. INTRODUCTION AND RECOMMENDATIONS

Research at the forefront of contemporary and future science and technology will demand adequate access to supercomputer power. There are already important problems within the Energy Research (ER) program that require major supercomputer resources for their effective solution. The small amount of Class VI (current supercomputer) time available to the ER community is totally inadequate. During FY84, when a small amount of such computing was made available to the ER community through the Energy Sciences Advanced Computation Program (ESACP), the hours requested exceeded the availability by an order of magnitude.

The disciplines with unsatisfied supercomputing needs include Nuclear and High Energy Physics, Chemical and Materials Sciences, Engineering and Applied Mathematical Sciences, Geological and Meteorological Sciences and the Biological and related sciences. Extensive computing requirements in these fields have already been identified, however, new problem areas are continually being uncovered and the magnitude of the latest demand for supercomputing in the ER programs is just beginning to be understood. However it is clear that the totality of the needs in these and related disciplines is much larger than is currently supported. Some of these needs are documented in the body of this report.

During part of FY83 and through all of FY84 the Office of Energy Research made available about 5% of the available hours on two CRAY 1 computers at the National Magnetic Fusion Energy Computer Center (NMFECC) to start serving at

53-831 O -86 - 10

least a part of these requirements. Furthermore a CRAY XMP will be acquired in FY85 on an interim basis, to start to serve more fully the Class VI needs of the ER community. In order to bring such a facility to a useful working level for the user community as soon as possible, this computer will be operated by the staff of the NMFECC, and integrated into the existing system.

The requirements for aggregate capacity and for capability for large problems demands the acquisition of yet more powerful computers than a single Class VI computer. The capability issue is primarily a matter of the limited addressability for a single program in today's Class VI computers. Some problems require memory sizes an order of magnitude larger than available in Class VI machines. Other capability limitations arise in problems where it is impractical to exploit in a rational way the high performance features of the current vector oriented supercomputers. Other architectures now under development may prove to be more effective for a wide range of ER problems.

This report was compiled by the ER Users Advisory Committee (ERUAC) and contains representative examples of known unsatisfied computing needs in DOE's Energy Research community. To accommodate these needs sensibly, the ERUAC makes the following recommendations:

RECOMMENDATIONS

1. The Office of Energy Research should request FY86 funds to acquire a Class VII mainframe to serve the supercomputing needs of the ER program on a national scale. The Class VII system should have computing speeds (performance) several times faster than current Class VI computers, and a

memory large enough so that a single program can address at least 32M words.

2. The details of the organizational structure that will deliver large- scale computing services to the ER community in the future should be given continuing review. Whether there is a single center built on the present NMFECC, a new single center, or several centers possibly built around alternative architectures, and whether the present NMFE network forms the basis of a future network depends on both technical and institutional issues that require additional study.

3. The interim strategy to acquire a Cray XMP early in FY85 to be integrated into the NMFECC is strongly supported. The Office of ER should proceed with all possible speed to consummate that plan.

4. Within the Energy Sciences Advanced Computation Program resources must be made available for the development and support of a capable communications network. Such a network should have adequate bandwidth to accommodate modern interactive graphics computing techniques with growth capabilities for future needs. To the largest extent possible it should provide access mechanisms (gateways) to local, national and international networks important to the ER community.

5. The existing NMFE communications network, including local equipment, should be extended as soon as possible to allow effective utilization by the ER user community of the interim Class VI computer that is to be installed at NMFECC.

6. As the Program develops and the magnitude of the supercomputer needs of the ER community becomes better understood, in both total capacity and in the capability that will be needed for specific problems, appropriate additional advanced (Class VII and beyond) supercomputers should be acquired to maintain U.S. energy research leadership. The first review point will come early in calendar 1985, at which time a preliminary decision will be made when the first additional machine must be procured. In Chapter 4 of this report, descriptions are given of a few important problems that are already known to require higher capacity and capability than Class VII computers.

2. ENERGY SCIENCES ADVANCED COMPUTATION

During FY84, the Department of Energy (DOE) created the Energy Sciences Advanced Computation activity and established, as its major program, a supercomputer access program. This program was initiated as the result of many panels and committees which had investigated the area of availability of modern supercomputer resources to the scientific research community within the U.S. and to the DOE research community in particular. It was found that the current availability of modern supercomputer resources within the U.S. falls far short of the amount of these resources needed by the research community and it was also found that modern supercomputers themselves do not have sufficient capability to address many of the computational needs of this community. During FY84 a requirement analysis was conducted throughout the research community which is funded by the Office of Energy Research (ER), and this analysis verified that several Class VI computer systems would be needed to begin satisfying this supressed demand. The purpose of the Energy Sciences Advanced Computation Supercomputer Access Program is to provide nationwide high-speed network access to modern centralized facilities within the constraints of budgetary resources.

In order to begin addressing this access problem as quickly and as economically as possible, ER decided to utilize the existing National Magnetic Fusion Energy Computer Center (NMFECC) and its installed high-speed satellite network, which is described below, across all ER programs. During FY84, ER allocated five percent of the NMFECC resources to the non-fusion ER programs. Because the NMFECC satellite network was already accessible at many DOE laboratories and universities and because this network provides gateways to

other networks (such as ARPANET, TYMNET, and in the future other networks of importance to the ER community), many researchers were able to gain access to the NMFECC facilities with very little lead time and minimal additional cost. As a result, this FY84 allocation was fully subscribed and utilized at a substantial savings over alternative methods of providing similar resources.

This Supercomputer Access program is managed as in other research activities. The resources are allocated throughout the research community on the basis of scientific merit and need. At the beginning of each fiscal year, requests for supercomputer resources are submitted to the ER program offices at DOE Headquarters. These requests are evaluated and ranked by the ER program managers. Resource allocations are then determined based on this ranking and on the total resources that will be available during a given fiscal year. These allocations can be adjusted, as necessary, during the year by requests to ER program offices and implemented by the NMFECC staff.

For fiscal year 1985, the Office of Energy Research is funding a Cray X-MP/2 computer system for installation at the NMFECC to further expand the availability of supercomputer resources to the non-fusion ER programs. This system will be operated on an interim basis to address the near term capability and capacity needs. The Office of Energy Research plans to replace this Cray X-MP/2 system with a more capable Class VII system in FY86 in order to provide the capabilities required for the computational needs described in this report. The Class VII system will be acquired through a competitive procurement at a time when U.S. vendors are expected to market at least three systems of this capability. The site at which this system will be located is, at this time, undecided. During FY85, the Energy Sciences Advanced Computation Staff will

analyze the future communications needs of this access program and the logistics and economics of establishing a permanent site or sites dedicated to this purpose.

New Supercomputers

Supercomputers are the most powerful general-purpose computers available for information processing. Currently, supercomputers have the capability of performing hundreds of millions of arithmetic or floating point operations per second (MFLOPS) and are used in two general areas: real-time applications such as signal processing and scientific computing. In the race to build the next generation of supercomputers, scientists are experimenting with a variety of architectural designs. The new architectures will have as few as two processors with shared memories to extensive parallel architectures with hundreds of local memories and processors, all executing instructions simultaneously.

During the past 35 years computers have grown more powerful with speeds that were unimaginable only a few years ago. The speed of computation achieved by the fastest computers has increased by a factor of 10 million from the 1940's to the present. However, the supercomputer industry is facing a demand for additional speed improvements of at least two hundred times in this decade. Unfortunately, the projected speed increases available from faster components appear to be limited to at most about a factor of ten during the next decade. Hence, it is clear that new forms of parallel computer organization will be required.

Architecture of New Supercomputers

There are three types of parallel architecture capable of increasing performance by a hundredfold over today's state-of-the-art supercomputers. They are:

- Lockstep vector processors,
- Tightly coupled parallel processors, and
- Massively parallel devices.

Concurrency is the avenue to increase speed with a given component technology. Today's supercomputers eliminate the overhead of instruction processing associated with each datum in conventional machines. They permit a single instruction to specify that an operation be performed on multiple operands and differ from conventional systems primarily in the high bandwidth transfer of multiple data elements to and from the execution unit under the control of a single instruction.

When the execution unit operates simultaneously (in lockstep) on many data entities, as in the ILLIAC-IV, the machine is said to have an array architecture. The ILLIAC-IV contained an 8 x 8 array of floating point processor elements (PEs). Keeping most of the PEs busy most of the time is the major software and algorithmic challenge.

When the execution unit operates on sets of data, on an assemply line basis, the machine is termed a vector processor or pipeline processor. The CDC 205 and Cray 1 are examples of vector processors. The real beneficiaries of such vector processors have turned out to be multi-dimensional fluid type codes, which are dominated by long vector loops. They run at astonishing speeds relative to more trivial, scalar bound, calculations.

A second architectural type capable of a hundredfold increase over state-of-the-art supercomputers is tightly coupled systems of a few high-performance processors. In principle, collaboration of these processors on a common task can produce the two orders of magnitude speedup that is needed. Admittedly, there are important architectural issues in tightly coupled systems, especially communication geometry between processors and memories. However, algorithmic issues for them are somewhat less complicated than for lockstep vector processors.

The current trend in supercomputer architecture is toward tightly coupled systems with two to four vector processors typically sharing a large memory. Recent experiments suggest that these systems can be successfully used in parallel processing of scientific computations.

The next logical step in this trend is toward systems with 8, 16, or more processors. Whereas the scientist may successfully find sufficient concurrent tasks to achieve efficient use of a system of four processors, success on systems with 64 processors, for example, is by no means assured.

In the long term it may be possible to build massively parallel systems, that is, systems with 1000 or more processors communicating with thousands of memories. In general, the scientist cannot manually find and manage parallelism for thousands of processors. Rather, the software must find it, map it onto the architecture, and manage it. Therein lies a formidable research issue for massively parallel computation.

The following two tables list 1) existing supercomputers, and 2) announced supercomputers. This tabulation employs only the few parameters usually contained in press-release-type information.

Table 1 shows supercomputer systems currently being installed. The first of the new generation of supercomputers, the CRAY 1S, is no longer being manufactured. It is not shown in the table although over fifty of them have been installed since 1976. It has been replaced by the CRAY XMP. Approximately 15 CYBER 205s have been delivered. The first Hitachi S-810/20 has been delivered to the University of Tokyo. The first Fujitsu VP100 (a smaller version of the VP200) was delivered to the University of Nagoya in November 1983. It is to be used primarily for plasma physics research.

TABLE 1

Current Supercomputers

Organization	Fujitsu	Hitachi	CDC	CRAY	CRAY	CRAY
Model	VP-200	S-810/20	205	X-MP/1	X-MP/2	X-MP/4
announcement	July 1982	Aug. 1982	June 1981	Aug.1984	Aug. 1982	Aug. 1984
architecture (64 bit words)	vector (IBM compatible)	vector (IBM compatible)	vector	vector 1 CPU	vector multi-processor 2 CPU	vector multi-processor 4 CPU
maximum performance (M FLOPS)	500	630	400	252	479	953
maximum main memory size 64 bit words	32M MOS	32M MOS	16M MOS	4M MOS	4M Bipolar	8M Bipolar

Table 2 shows new systems in design or early stages of prototype development. NEC is the latest to announce its entry, the SX 2, for delivery in 1985. The Cray 2 is the most developed and it is expected that several may be delivered during 1985. However, it seems to be a prototype system for the newly designed Cray 3. The latter is intended to be the first Gallium Arsenide based supercomputer if this technology proves successful. The Cray 3 , ETA GF 10, and Denelcor HEP 2 are promised for delivery in 1986.

TABLE 2

Supercomputers Now In Design

Organization	CRAY	CRAY	ETA	Denelcor	NEC
Model	2	3	GF10	HEP-2	SX-2
announcement (or project start)	none officially	none officially	Sept. 1983	May 1983	April 1983
availability	1985	1986	1986	1986	1985
architecture	vector multi-processor 4 CPU	vector multi-processor 8 CPU	vector multi-processor 8 CPU	scalar multi-processor 64 CPU	vector
maximum performance (M FLOPS)	1,000	not available	10,000	4,000	1,300
maximum main memory size 64 bit words	256M MOS	not available MOS	256M MOS	256M MOS	32M MOS

National MFE Computer Network

The purpose of the MFE Computer Network is to provide to fusion researchers in the U.S. the full range of available computational power in the most efficient and cost effective manner. This is achieved by using a network of computers of different capability tied together and to the users via dedicated data lines and dial up telephone lines. The existence of this nationwide computer network allows projects to be sited anywhere in the country without regard to local computer availability, and therefore increases enormously the flexibility of the fusion program.

Levels of Computer Capability in NMFECC

The concept of the NMFECC is that different levels of computer capability are provided at the various locations according to research priorities. At the national center, providing high level capability to the entire community, is a CDC 7600 and two high-speed CRAY 1 computers with one and two million words of memory, respectively. Additional equipment at the national center includes processors and other ADP equipment for communications, file management, and data storage. During fiscal year 1984, five percent (5%) of the capacity at the NMFECC was allocated to the Energy Research programs other than Magnetic Fusion.

In FY1985, there are planned acquisitions for supercomputers for both the Magnetic Fusion program and the other Energy Research programs. A CRAY X-MP/2 is planned for October 1984 installation to provide additional support for the other Energy Research programs and a CRAY-2 computer system is planned for

April 1985 installation to support primarily the Magnetic Fusion requirements. (Figure 1)

At the next level of capability are User Service Centers (USC's): DEC-10 computer systems with direct high-speed access to the national center through PDP-11/40 remote communications control processors. There are now five operational USCs (Figure 2) in the field located at Princeton Plasma Physics Laboratory (PPPL), the Los Alamos National Laboratory (LANL), the Oak Ridge National Laboratory (ORNL), GA Technologies, Inc. (GA), and LLNL (for the mirror confinement program). A sixth USC, used in center operations, is located at the NMFECC itself.

A third level of capability is provided through the Network Access Port (NAP). MFECC designed the NAP to permit remote computers to be connected to the MFE network as remote hosts. There are currently five NAPS installed, all of them connecting VAX 11 series computers into the network.

A fourth level of capability is provided by Remote User Service Stations (RUSS) at selected sites. (Figure 2) RUSS stations provide users with the capability of printing output files locally on a 1000 line/minute printer and act as a terminal concentrator for up to 16 interactive terminal users. RUSS stations are connected to the nearest MFE-NETWORK communications processor over a 9600 baud dedicated line. (Figure 2)

<u>Data-Communications Systems</u>

Data Communications services to the National MFE Computer Center are provided on a 24 hours/7 day basis. Three types of service are provided to NMFECC users as outlined below:

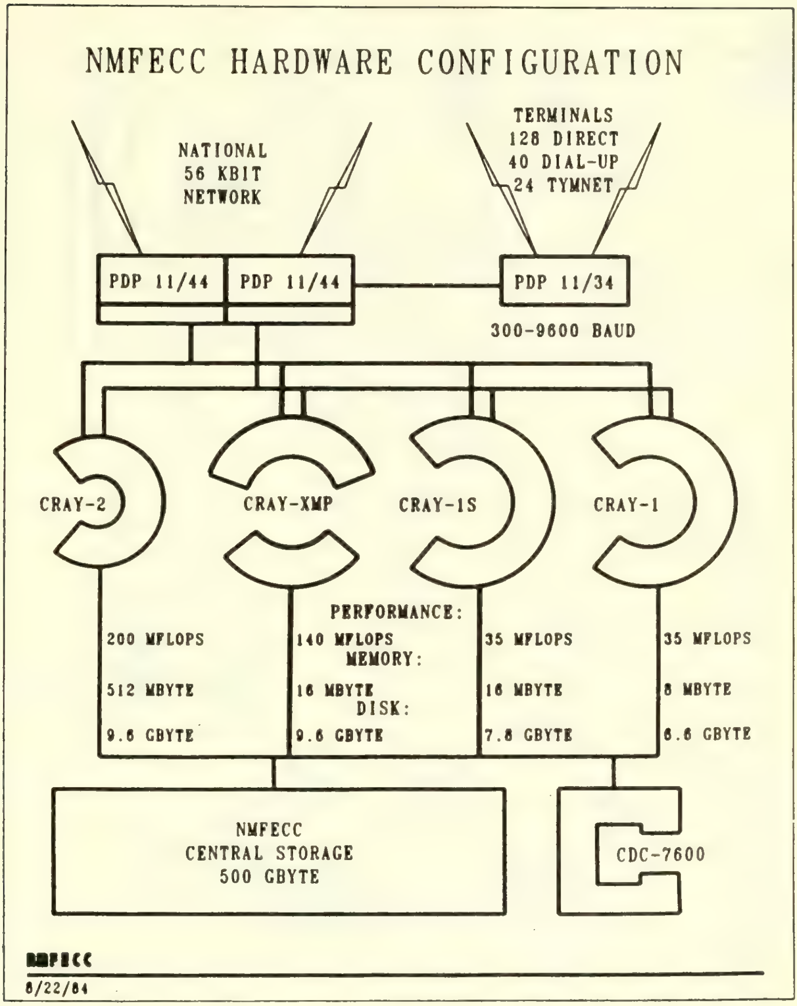
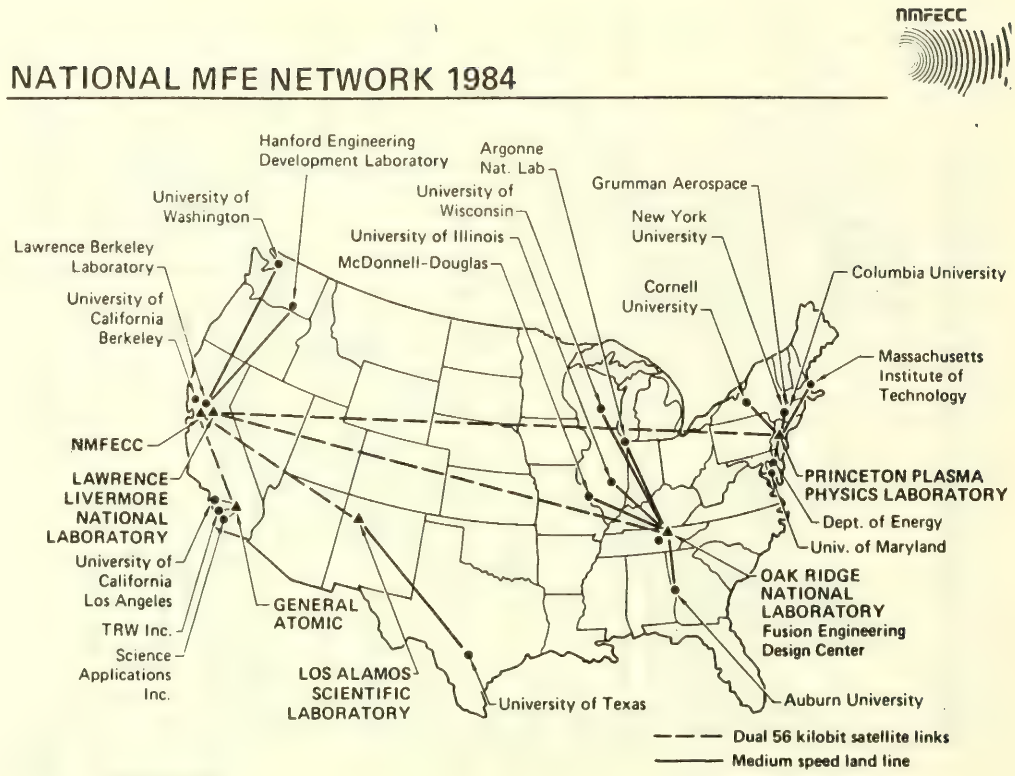

NMFECC HARDWARE CONFIGURATION
NATIONAL
56 KBIT
NETWORK
TERMINALS
128 DIRECT
40 DIAL-UP
24 TYMNET
PDP 11/44
PDP 11/44
PDP 11/34
300-9600 BAUD
CRAY-2
CRAY-XMP
CRAY-1S
CRAY-1
PERFORMANCE:
200 MFLOPS
140 MFLOPS
35 MFLOPS
35 MFLOPS
MEMORY:
512 MBYTE
16 MBYTE
16 MBYTE
8 MBYTE
DISK:
9.6 GBYTE
9.6 GBYTE
7.8 GBYTE
6.6 GBYTE
NMFECC
CENTRAL STORAGE
500 GBYTE
CDC-7600
NMFECC
8/22/84

Figure I.



NATIONAL MFE NETWORK 1984
NMFECC
Hanford Engineering Development Laboratory
Argonne Nat. Lab
University of Washington
University of Wisconsin
Grumman Aerospace
New York University
Lawrence Berkeley Laboratory
University of Illinois
McDonnell-Douglas
Columbia University
Cornell University
University of California Berkeley
Massachusetts Institute of Technology
NMFECC
LAWRENCE LIVERMORE NATIONAL LABORATORY
PRINCETON PLASMA PHYSICS LABORATORY
Dept. of Energy
Univ. of Maryland
University of California Los Angeles
OAK RIDGE NATIONAL LABORATORY
Fusion Engineering Design Center
GENERAL ATOMIC
TRW Inc.
Science Applications Inc.
LOS ALAMOS SCIENTIFIC LABORATORY
University of Texas
Auburn University
Dual 56 kilobit satellite links
Medium speed land line

Figure 2

1. Wide band (56 KB/sec) Satellite Network Service. Users at Major USC's on the MFE net may log on to their local DEC-10 system and interact with the computing resources at the Central facility in Livermore.

2. Dedicated 9600 Baud Service. Remote User Service Stations on the MFE Net are served by dedicated leased 9600 baud lines which terminate either at the Center (LLNL) or at the nearest MFE Communications Control Processor. (Figure 2)

3. Dial Up Service. Users not at major fusion laboratories may dial-up the Center using one of the following services: (a) TYMNET, (b) ARPANET and (c) DIRECT DIAL COMMERCIAL.

NMFECC Computing Environment

The NMFECC computing environment reflects the needs of computer users in the Magnetic Fusion Energy research community. Both interactive timesharing and batch processing are available. The fusion community has always found that interactive computing, even with the largest codes, is by far the most efficient use of physicists efforts. The 5% overhead in swapping codes in and out of the machines provides fast debugging, immediate turn around on key results, and the capability to interact with codes which need user control. The Livermore Time Sharing System (LTSS) was adapted by the NMFECC for the CRAY 1 computer in about six months. CTSS is supported by libraries of FORTRAN callable subroutines which enable a user to issue almost every system call,

giving access to every part of the hardware. A typical physics code can be run from a terminal, display graphics as it runs, be interrupted or interrogated at any time. The ability to start or stop a code at any point and inspect the results provides debugging at least 100 times faster than older methods. The CTSS operating system will also be used on the CRAY X-MP/2 and the CRAY 2.

NMFECC has designed a multi-level file storage system called FILEM. FILEM is a highly versatile system which allows users to store and retrieve programs and data files in the central computing facility at Livermore for an indefinite period of time. FILEM has been designed to accommodate the needs of users at remote sites. The CDC 7600 has been programmed to assume virtually all of the tasks associated with file custodianship including indexing, storage, retrieval and efficient management of the file storage media.

It is the policy of NMFECC to make all computer documentation available on line so that users may provide themselves with up-to-date system documentation by simply printing out the document at their local printer or terminal. Any part of any document may be displayed on remote terminals and the routine DOCUMENT is capable of scanning text for the user to locate a specific topic of interest. NMFECC has provided the user with two routines called TELL and NEWS which allow users to send a message or question to any other user on the NMFECC network. NEWS and TELL are also commonly used by users to ask NMFECC staff about specific problems they have encountered. NMFECC systems programmers and documentarians use NEWS to broadcast all system or documentation changes.

3. ENERGY RESEARCH SUPERCOMPUTING PLANNING PROCESS

The Supercomputer Access Program of the Energy Sciences Advanced Computation activity is managed by the Scientific Computing Staff of the Office of Energy Research (ER) within the Department of Energy (DOE). This program is managed within the framework of the DOE 5 year ADP planning process as described in the DOE ADP Long Range Plan. During each fiscal year supercomputer requirements are assessed across all ER programs. This is accomplished via a "call" issued by the Scientific Computing Staff and other ER program managers to ER funded researchers. This call requests submissions for supercomputer access, including the amount of computer time needed and a justification for this need.

These requests are used for two purposes. First, an ER wide computer resources committee will analyze these requests, review the recommendation from each program office, and will make the time allocations for supercomputer resources for the next fiscal year. Requests and allocation modifications which are made out-of-cycle will be evaluated in the same manner.

The second purpose of these requests is to provide planning information for the DOE ADP long range planning process. This process, identifies major ADP resource requests within the DOE, validates the programmatic requirement, encourages free marketplace competition, and seeks to achieve the most cost effective and efficient operation of ADP resources throughout the DOE.

The need for more advanced computers in 1985, 1986, and 1987 should properly be discussed from two perspectives. The first persepctive has to do with capacity and suggests the question: With a CRAY XMP in operation does

NMFECC still fall short of the projected requirements of the ER program? The second perspective has to do with capability and can be discussed in the context of the following question: Are there important ER problems that cannot be solved with todays most capable computers? The answer to both questions is yes.

In order to assess the implications of new computer acquisitions, we discuss the concepts of computer capability (power) and capacity.

Capability (power) is a measure of the size or complexity of a job a computer system can process within some specified time. In practice, this period of time can be from several seconds to several days.

An increase in capability of a computing system implies an increase in functions or the ability to solve more difficult problems: problems requiring computer models with higher dimensionality, finer resolution, more physical effects, or other improvements which tax computing power.

Capacity is a measure of the throughput of a computer system, i.e., the number of jobs of a given type that can be accommodated within some specified time.

An increase in capacity implies the ability to do more of the same type of job. The capacity of a computing facility may be increased by the addition of similar computers, or by the provision of computers with greater capability.

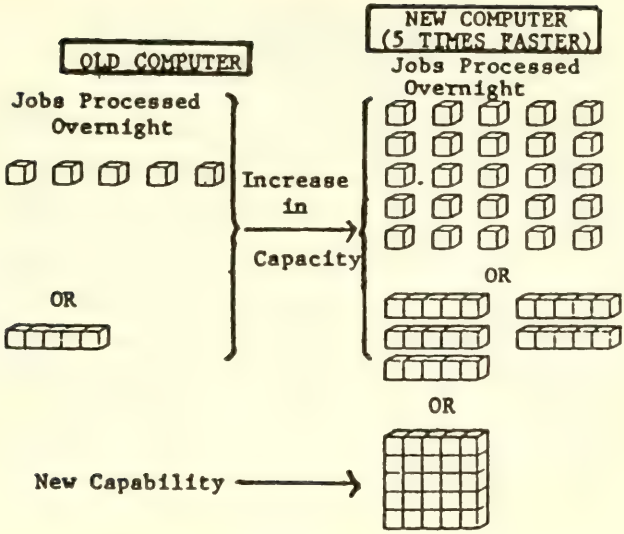


Figure 3. A faster computer provides additional capacity and capability (power). Both are important. However, most often, it is increased capability--the ability to tackle harder problems with better tools--that provides the major payoff. Much of this report focuses on the implications of acquiring new computer capabilities.

Capacity

During FY84 the requests for supercomputer access by the various Energy Research programs were analyzed and validated in order to make allocations for FY85 and to project the need for the outyears, particularly the need for the Planning Year FY86. The total requests for access for FY85 by the High Energy and Nuclear Physics, Basic Energy Sciences, Applied Mathematical Sciences, and Health and Environmental Research Programs exceeded the capacity of three (3) Class VI Computer Systems. This requirement in NSU's, versus the planned, installed capacity to support this requirement, is shown in Figure 4.

	FY84	FY85	FY86	FY87
Required	57,700	140,100	290,700	603,000
Installed	4,000	68,000	140,000	250,000

The requirement for supercomputer access to support the large scale computational needs of the researchers in these Energy Research programs far exceeds the currently planned installed capacity for the foreseeable future. Indeed, the Class VII computer system which is planned to be installed to provide additional capability as described in the next section, will be over-subscribed upon installation.

The supercomputer requirement broken out by program within Energy Research is presented in Table 3.

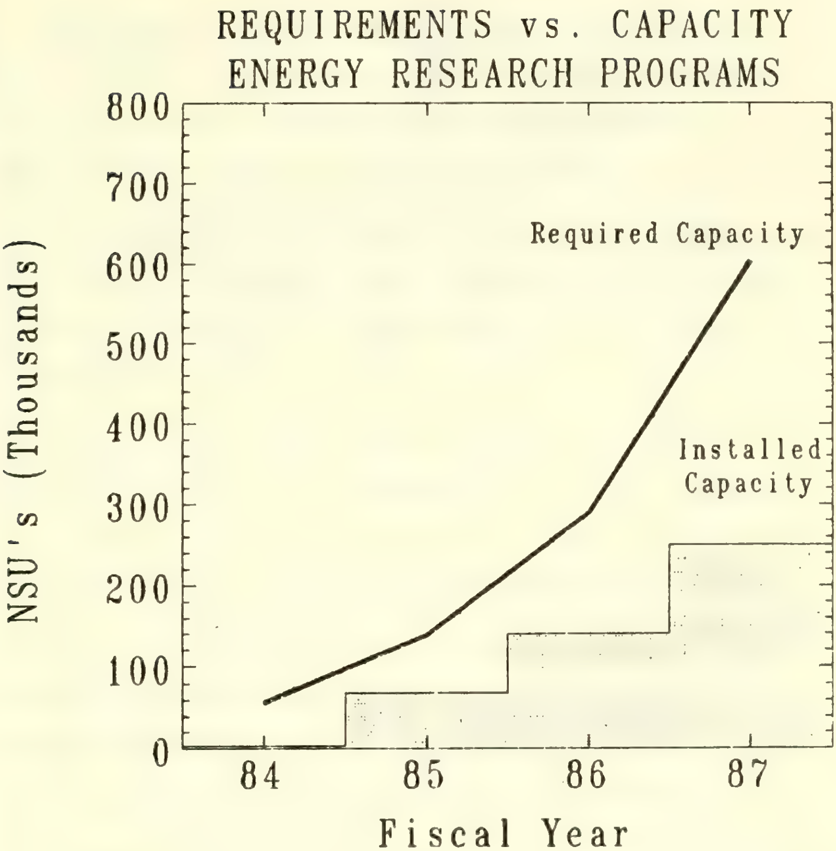

REQUIREMENTS vs. CAPACITY
ENERGY RESEARCH PROGRAMS
NSU's (Thousands)
800
700
600
500
400
300
200
100
0
Required Capacity
Installed Capacity
84
85
86
87
Fiscal Year

Figure 4

TABLE 3

Energy Research Supercomputer Requirements in NSU's

	FY84	FY85	FY86
HENP	25,400	62,600	157,900
BES	25,100	52,900	85,800
AMS	4,200	17,300	34,800
HER	3,000	7,300	12,200
Totals	57,700	140,100	290,700

Capability

Historically, scientists who use supercomputers have constrained their numerical simulations to an average execution time of about ten hours. This constraint reflects the scientist's need to make daily progress. Thus, the amount of complexity incorporated in models is scaled to the computer's ability to produce results in about a ten-hour execution time. The capability of a supercomputer dictates the amount of complexity that can be treated. Because of this limitation, scientists engaged in large-scale numerical simulation have continually sought bigger and faster computers. Today, scientists engaged in energy research need supercomputers that are up to 200 times faster than state-of-the-art equipment.

In order to understand the requirements for more powerful computers, we must explore the generic reasons for having increased computational speed and storage.

Dimensionality. The real world exists in three space dimensions plus time. If computational models reflected the real world exactly and completely, they would treat all four of these dimensions and other parameters that are equivalent to additional dimensions. With current computers, it is possible to treat two space dimensions and time for some problem types, three space dimensions for others, and three space dimensions plus time for a very limited set of problems. Speed increases of about a factor of 200 in this decade are needed to allow researchers to solve urgent multidimensional problems that are now intractable.

Resolution. Every region of space contains infinitely many points. Thus, the first step in modeling any natural phenomenon is to approximate the space with a finite set of zones, each of which requires a number of calculations. Increasing the number of zones means we can determine more completely and accurately what is happening in any environment, but the computational time grows very rapidly. For example, in a two-dimensional time-dependent model, the running time grows in proportion to the third power of the increase in resolution; increasing the number of zones by just a factor of 2 would increase the time to complete the problem by a factor of 8. Many complex problems now run up to 100 hours, so it is clear that resolution increases of even relatively small factors can overwhelm the capabilities of current supercomputers.

Physics. All computational models dealing with the frontiers of science and technology make simplifying assumptions about the laws of physics in order to keep the calculations from running too long. In some models, including just one additional physical effect can increase running time by a factor of 10.

Faster supercomputers with much larger memories will permit researchers to solve problems that cannot now be economically solved.

Combination effects. Although dimensionality, resolution, and physics each have powerful effects on running time by themselves, the overall needs are derived from combinations of these effects. The highly complex problems now being studied in energy research programs, require computational models with higher dimensionality, and with higher resolution, and with more physics.

4. THE ROLE OF COMPUTER MODELS IN ER PROGRAMS AND THE NEED FOR MORE POWERFUL COMPUTERS

We can consider areas of computing of importance to the Energy Research program. As the fields of physics advance, theoretical investigations naturally evolve from a stage where the most important problems can be solved analytically, to where numerical solutions become essential. At any one time, different sub-fields are in different stages of this process. However one can identify—even across the boundaries of different subfields--common features that drive this evolution. In general, large-scale computation becomes necessary as the systems of interest become more complex. One moves from a few degrees of freedom to many degrees of freedom.

Complexity emerges in different ways in the different subfields of physics. In some cases, it is in moving from ordinary differential equations to partial differential equations for the solution of interesting problems. This has happened, long since, in fluid dynamics and for the Schrodinger equation in non-central or many-body conditions, for radiative transfer problems in astrophysics, and in other areas. Alternatively, in areas such as nuclear reaction theory, many coupled integro-differential equations naturally arise and must be solved self-consistently.

Complexity also increases dramatically when one moves from simple or one-dimensional models of physical processes, to realistic simulations. One sees this in research on the phases of fluids and solids and lattice models, in

plasma simulations, in many-body calculations of galactic dynamics, and in quantum field theory.

Complexity can come from the impetus to move from low order to high order expansions. This is the case in quantum electrodynamics, quantum chromodynamics; in high temperature and other expansions for liquids, solids, and lattice systems; and in high partial wave expansions for nuclear reactions. When the expansion techniques are algebraic, one moves from hand-algebra to algebraic symbol-manipulation on the computer.

Complexity arises in moving from scalar systems to vector or tensor systems, and from linear systems to nonlinear systems. A striking example of this is general relativity, whose partial differential equations are both tensor and highly nonlinear. The fluid dynamics of many-component systems is another example.

As one surveys some subfields of physics, as we shall now do, one finds over and over again that (i) there are important problems whose solutions must be found by computational techniques, and (ii) that even the best investigators are resource-starved, without the facilities for accomplishing even the tasks already at hand.

4.1 HIGH ENERGY AND NUCLEAR PHYSICS

Introduction

The requirement for computers capable of meeting the data reduction needs of a high energy physics laboratory has, historically, been so great that all other computing requirements could be met without significantly impacting the large central facility. However, in the decade of the '80's, several new computational needs have appeared which require the unique capabilities of the Class VI Super Computers and, in a few cases, clearly require capabilities presently associated with Class VII systems.

The theoretical high energy physics community represents an important class of users with very large unsatisfied computational needs. This is primarily due to the rapid rise of computational quantum field theory, particularly in numerical studies of lattice gauge theory. To put this development in perspective, it should be noted that computer simulation is a generic numerical tool for studying the behavior of particles and fields, and its importance does not rest on any particular fashion nor on the currency of any particular theoretical idea. The ability to carry out such calculations is primarily a result of the rapid increase in available computer power, and as such, it represents a permanent change in the way theoretical physics is done. The needs here fall into two distinct categories. The first category includes the more traditional forms of theoretical computation such as numerical integration, solution of integral or differential equations, calculation of Feynman diagrams, etc. The second category is the large scale numerical simulation of quantum field theory on a lattice. These calculations are highly

CPU intensive. The lattice gauge theory algorithms are relatively simple, repetitive and easily vectorizable. Thus, they are well suited to a variety of parallel and pipelined architectures provided that a large, faster accessed memory is available. Even low statistics calculations on modest sized lattices require the equivalent of tens of CRAY hours.

Two newly emerging needs for computer power beyond the scope of Class VI systems are from the accelerator and experiment design communities. An example of an accelerator design requirement is for the turn-by-turn simulation of potential designs for the new superconducting super collider (SSC) accelerator currently in conceptual design. The integrated time needs here are CPU times measured in CRAY-1 equivalent years.

An example of the experiment-design-related requirement is the full simulation of Monte Carlo events in a colliding beam detector system. The number of simulated events run should, ideally, be substantially greater than the number of real physics events to be analyzed. Furthermore, since experimental results may change the way a detector is tuned, it may be necessary to make the simulations concurrently with the taking of data, i.e., when the data reduction computers are most fully loaded.

Experimental high energy physics data reduction, which has heretofore used standard general purpose computers, also needs a new genertion of computers. The generation of detectors now just coming into use necessarily gather data at very high rates in order to extract the physics of interest from the enormously large accompanying backgrounds. The volume of data collected from these new experiments is several orders of magnitude larger than in experiments performed

in the 1980 period. The computational problems are enormous and new classes of supercomputers along with special purpose processors appear to be the only practical way in which to satisfy these unfilled computational needs.

Accelerator Physics Simulations

There are two general types of accelerator simulations, both requiring large amounts of computation, that are particularly suited to the use of supercomputers. One is the simulation of high intensity beams where the self fields of the particles acting on each other is important. This type of simulation is in the same family as the plasma physics studies which are a large part of the work presently being done at the NMFECC. Thus the same, or similar, computer codes can be used. Examples are: high intensity accelerators, including all types of accelerator injectors, radio frequency (rf) sources (klystrons, gyrotrons, etc.) and a number of new accelerator and rf source ideas. Accelerator schemes using collective effects, laser accelerators, and free electron lasers (FEL) all fall in this category. An example of this type of calculation for a "Lasertron", which is a new high power rf source, is discussed later in this section.

The second general type of accelerator simulation that requires the power of supercomputers is the simulation of a long beam transport line. The design of the Stanford Linear Collider (SLC) or of a large circular storage ring requires establishing alignment and fabrication tolerances for dozens of different components. Since many copies of each component are required, the actual parameters can be assumed to lie randomly within the established tolerances. The accelerator designer "builds" a machine using a randomly

chosen set of all of these parameters and tests it for stability. Since a beam may have to circulate many cycles before encountering a fatal instability, the total computation time for each randomly chosen machine can be quite large. Also, because it is required to show that it is reasonably likely that a given actual machine will work, a large number of such random accelerators need to be simulated. Finally, because the cost of the accelerator is a function of the precision of the manufacturing tolerance, it is important to explore all implications of relaxing these tolerances. Two examples of this type of simulation for present day accelerators with currently available computers are given below. Even Class VI computers are not capable of adequately simulating a large storage ring such as the Superconducting Super Collider (SSC) which is currently in a planning stage for construction later in this decade.

An existing accelerator, such as the Tevatron at Fermilab, consists of approximately 200 quadrupoles, 800 dipoles and 200 correction elements. The projected SSC will have approximately 5 times this number of components. The anticipated behavior of an accelerator is predicted by using Lorentz's equations and tracing a particle through the magnetic structure of the computer model. A code, "TEVLAT", has been developed for this purpose at Fermilab. The measured magnetic properties of each of the magnets is used in the simulation. For magnets used in the Tevatron the parametrization of the magnet field requires 28 coefficients. (Address space limitations on the Fermilab computer have allowed only 12 of the measured 28 coefficients to be used.)

In order to study the behavior of the beam during, for example, the 30 second flat top ordinarily used for the fixed target at Fermilab, approximately 30 hours of Cyber 175 time would be needed for each condition studied in the

simulation. This is clearly impossible so the behavior of the beam is inferred from simulations corresponding to ~ 2 μsec of actual beam time. This, of course, leads to serious questions about the adequacy of the simulations. With an existing accelerator the model can be tuned so that accurate results can be obtained and predictions for accelerator behavior can be checked using the accelerator.

The simulation of a projected accelerator design clearly cannot rely on modeling which must be tuned before the results are reliable. The model must also be capable of simulating the behavior of the beam for longer periods of time. The ratio of the time to simulate a revolution of the beam on the computer to the actual time the beam takes to go around the accelerator (~ 3.5×10^3) will at best remain constant as the size of the accelerator increases. Thus there is clearly a need for computers with substantially faster processors in order to perform efficient simulations of projected accelerators.

Simulations of Errors in the SLC Arcs

The Stanford Linear Collider, (SLC) has two long curved beam transport lines to conduct the intense bunches of 50 GeV electrons and positrons, respectively, from the SLAC Linac to the collision point in the SLD detector. At the detector, the two beams are focused to one micron in radius and are brought into collision. The magnet quality and alignment tolerances that will permit this performance are very tight.

Two different programs are in use to simulate magnet misalignments and fabrication errors and to estimate their influence on beam quality at the interaction point. Each of the many different runs corresponds to a different machine configuration due to the introduction of a large set of random numbers which in turn generate magnet and alignment errors. The two programs serve as checks on each other.

After predicting the errors, the programs are used to introduce corrections and predict the corrected orbit. Figure 5 shows the results of a typical uncorrected run (dots) and the corrected orbit (line) for one beam. Using the SLAC IBM-3081, 600 seconds, of CPU time are required to simulate a single new configuraion, and to introduce the corrections needed to achieve the required performance. During the past year this tolerance study has been the largest single user of computer time for accelerator physics at SLAC.

Computer Simulation of the Lasertron

The concept of the Lasertron can be most easily described by referring to Figures 6 and 7. A short bunch of electrons (short in time compared to the period of an rf cycle), is emitted from the cathode and accelerated by a dc potential toward the anode. An rf cavity, located just past the anode aperture is the output cavity. The rf frequency is determined by the pulse rate of the electron bunches. Although other types of solid-state cathodes are possible, if a photocathode is used, then the light source will likely be a fast pulsing laser, hence the name "Lasertron".

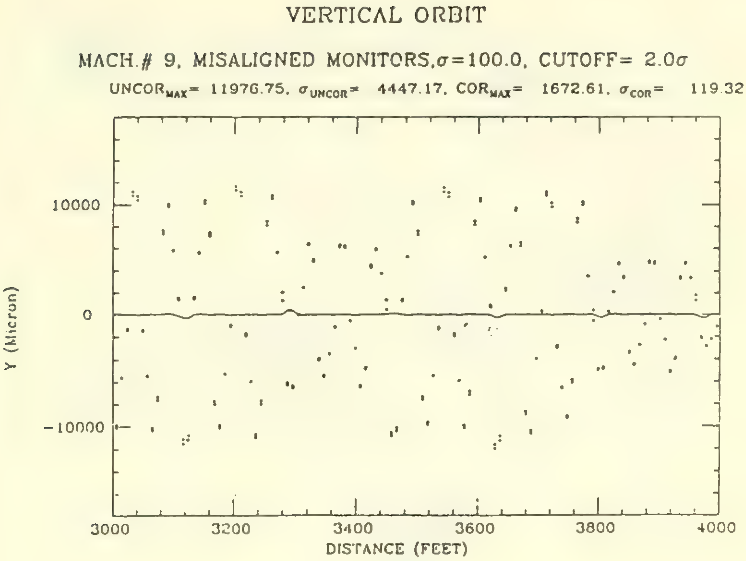


Figure 5 Beam orbit displacement is shown for an uncorrected beam line (dots) and for the corrected orbit (solid line) for one SLC arc.

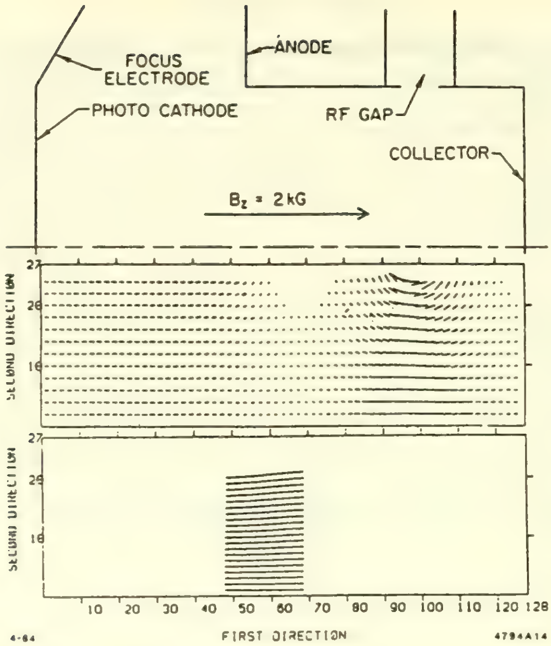


Figure 6. The layout of the complete Lasertron is shown in the top frame. The middle frame shows the rf field patters and, on the left, the dc diodie field patters. The lower frame shows a pulse of electrons moving from left to right.

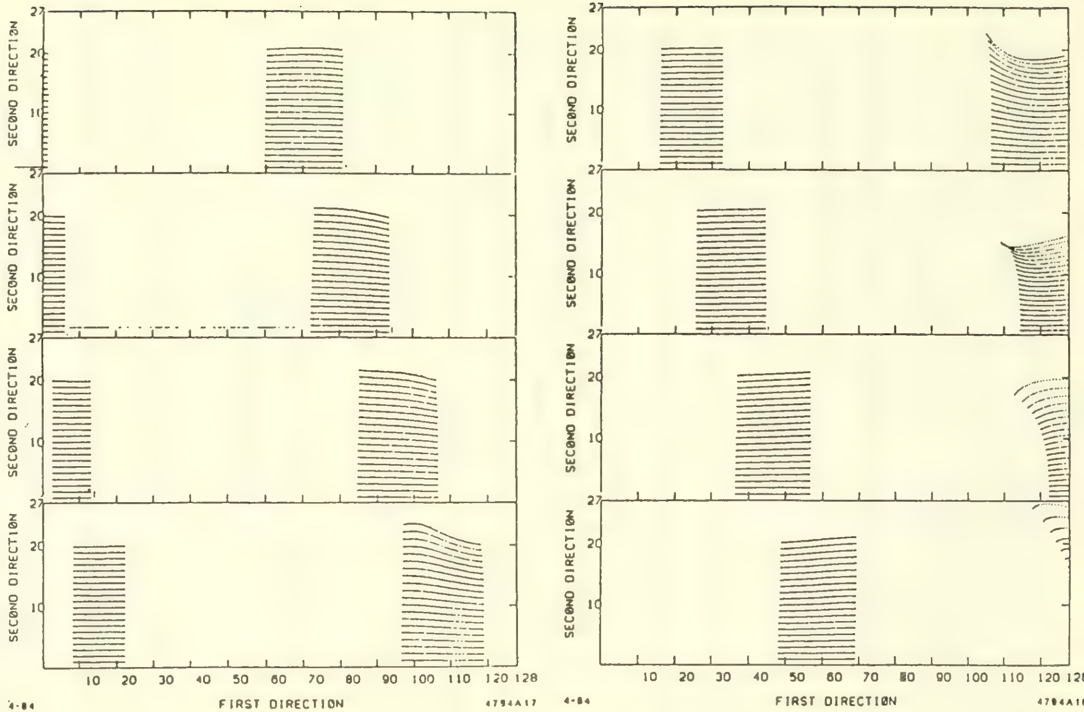


Figure 7. The Lasertron pulse is shown in eight sequential views during one rf cycle.

The Lasertron is simulated with the particle-in-cell plasma simulation program MASK running on the SLAC IBM 3081. MASK is also an extensively used program on the NMFECC CRAY system and is becoming an important tool for accelerator and rf source design studies. Because MASK runs about 10 times faster on the CRAY computers than it does on the IBM 3081, it will be possible to make much more detailed studies on the CRAY than the very preliminary results discussed below.

A complete Lasertron rf source, is shown in the Figures 6 and 7. An rf field has been included at a port part-way down the drift tube. The rf voltage at the port was 450 kV and the phase was adjusted for maximum energy extraction from the electron beam. For this case, the MASK diagnostics show that 28% of the initial beam energy per pulse is finally absorbed on the walls of the drift tube. The remaining 72% is thus extracted as rf energy at the port. This was for a beam current of 105 A peak, or an average beam power of 8.75 MW. The high output efficiency is an encouraging sign that the Lasertron could be a very efficient generator of rf power for future large accelerator projects.

Structural and Magnetic Field Calculation for Large High Energy Physics Equipment

The large magnet and detector structures needed for high energy physics detector systems are complex and very costly (millions of dollars). To minimize the programmatic and financial risks involved in the building of these systems great care must be taken in their design. Consequently, elaborate finite element analysis engineering studies must be made for structural elements before the design is completed and fabrication is started. For

magnets further magnetic field analysis are required, both from a structural stress point of view, and also for an analysis of the resulting magnetic fields, which the particles detected will traverse.

A number of general purpose programs are used today to analyze these systems. For the finite element analyses, there exist a number of commercially available programs. The program most widely used today in high energy physics for finite element analysis is ANSYS. The analysis of magnetic field shapes and the forces induced by these fields are supported by a number of different computer programs. Each of these extant programs is best suited to different classes of problems.

Some examples of problems of this class that have been computed recently are given here. In almost all of the large cases the detail of the analysis was limited by either the capacity or the capability of the available computing facilities. In these cases and for many of the future designs, the availability of a Class VI computer would be appropriate to solve these problems effectively.

Computer Design of a Large Detector Magnet

Figures 8,9 and 10 show a POISSON modeling of the SLD. This is a large, laminated solenoid particle detector which will be installed at the final focus of the SLAC Linear Collider, SLC.

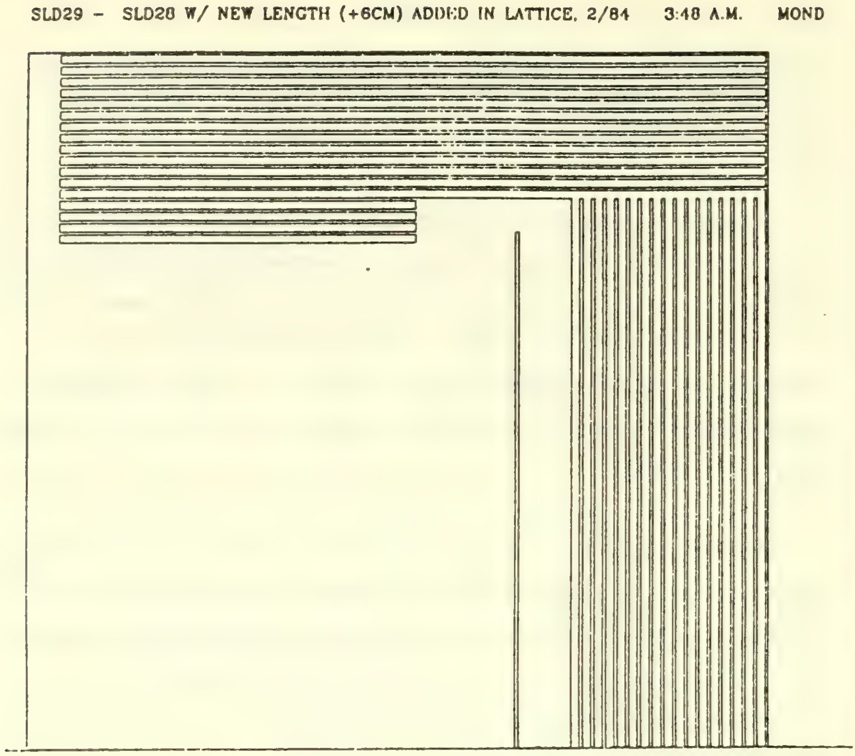


Figure 8. The physical layout of the laminated plates of iron for the SLD.

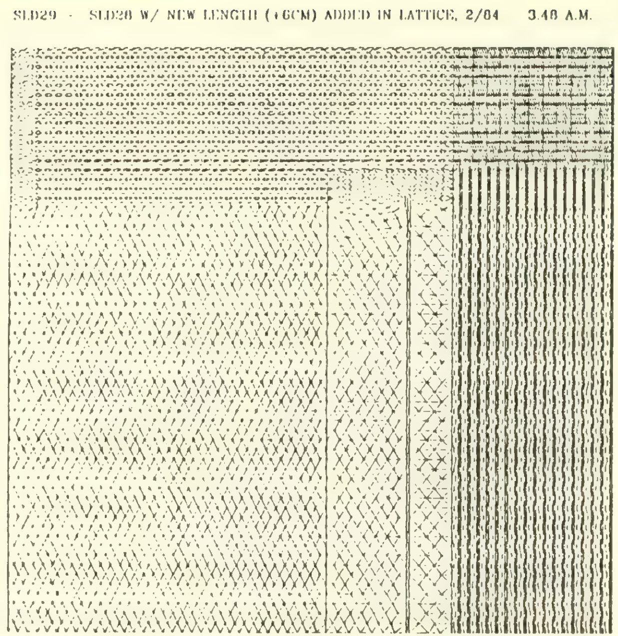


Fig. 9. The zoning of the triangular mesh for the SLD magnet.

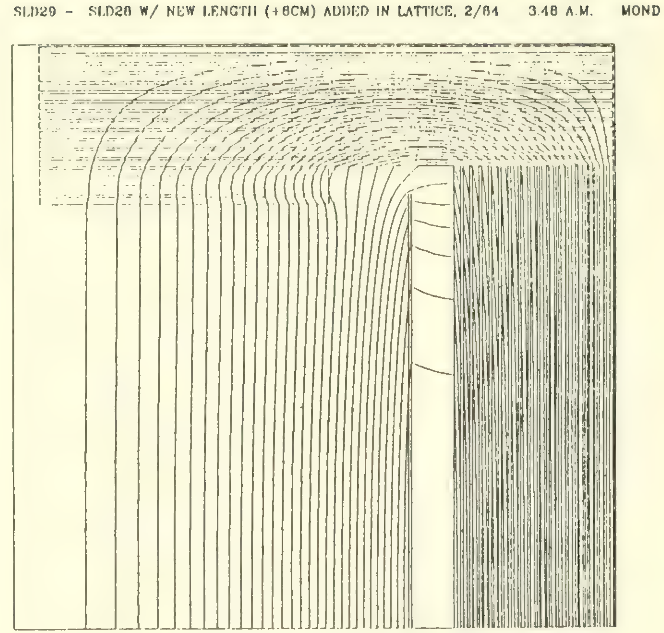


Fig. 10. Flux lines are superimposed on the layout of the SLD magnet

Both the end cap and the outer walls consist of 17 iron plates 5 cm thick and separated by 2 cm. Because of the large number of narrow, isolated iron regions, and the fact that certain areas of the problem are highly saturated, the problem is computationally unstable. In order to reach a solution, the problem must be linearized, and this in turn means that more iterations are required to reach convergence.

The running time for this problem is usually in excess of 45 minutes on the IBM 3081. This requires running in background which can take anywhere from 2 days (over a weekend) to 2 or 3 weeks. On the MFE CRAY 1S the running time for this problem is only 4 minutes.

The Fermilab Colliding Detector Facility

A large number of engineering studies have been made for the design of the Colliding Detector Facility (CDF) at Fermilab. These have consisted of finite element structural studies of the various detector chambers, calorimeters, magnetic yokes and coils and other ancillary structural elements. For most of these studies the program ANSYS was used. Typical runs for many of these studies ran over 1000 seconds on Cyber 175 computers. Ideally mesh sizes at least of factor of 2 finer would give a more refined and useful result. Since this would require more than 10 times the computing time, i.e. ~ 10,000 seconds, it was not practicle in the existing environment. Typically 15 or 20 such runs are required for each element so analyzed.

For the analyses of the superconducting coil for the detector, because of memory space limitations the coil had to be analyzed in parts. Typically this took 40 minutes of Cyber 175 time per analyses run. To analyze the complete coil in one unified run, with an appropriate finer grid, would take an estimated 1000 minutes per run. The availability of a Class VI machine would have made this a possibility.

Lattice Gauge Quantum Chromodynamics Calculations

Monte Carlo simulations of lattice gauge theories, and more specifically of Quantum Chromo-dynamics (QCD), while not being the only calculations of interest of particle theory, are presently the most demanding in computational resources and the most likely to produce quantitative predictions. Two essential elements enter into these calculations:

a) the generation of gauge field configurations distributed according to the exp{-S} measure;

b) the evaluation of quark propagators in the background of the gauge fields provided by the above configurations.

The degrees of freedom are made discrete by introduction of a (usually) hypercubical lattice. A lattice extending for n_s sites in the spatial directions and n_t sites in the temporal one entails $4n_s^3n_t$ gauge dynamical variables associated with the links of the lattice. At present, we lack a quantitative understanding of any collective excitation which may dominate the functional integrals. Therefore all of the above link variables must be treated on the same footing. The lattice must be sufficiently big to contain a hadron, and provide enough resolution so that a lattice with the same physical

volume but a finer subdivision would lead to essentially unmodified results (notion of scaling toward the continuum limit). Let us assume, to fix ideas, that the lattice extends for 10 sites in all spatial directions and 20 in the temporal one. This give a total of 80,000 link variables, i.e., 80,000 SU(3) matrices which must be kept in the memory of the computer for a simulation of QCD. Although SU(3) has only 8 independent parameters, recording the link variables other than in full matrix notation increases substantially the complexity of the calculation. Thus a lattice configuration corresponds to 80,000 x 18 real numbers = 1,440,000 words of memory. To proceed from one configuration to the next all 80,000 link gauge variables must be "upgraded". The upgrading of a single variable involves on the order of 4000 element arithmetic operations. We thus obtain an operation count of $\approx$ 320 million to generate a new configuration. Typically hundreds or thousands of configurations must be generated to reduce meaningful results. The calculations of the quark propagators are about as demanding in computational resources. Therefore, the following requirements emerge for this kind of numerical investigations:

i) more than a million words of memory must be accessible,

ii) sustained calculational speeds of several million floating point operations per second must be possible.

On point i): with a memory of 2 M words a simulation on a 10^3 x 20 lattice can be done working with fast memory only. With a smaller memory, or for large lattice size, schemes by which the variables are input from disk, evaluated and output onto disk again must be devised. This is not impossible, given the highly structured form of the calculations. For instance, the variables with a

definite time coordinate, together with their neighbors in time, can be brought into fast memory and upgraded; then those at the next time etc. (time slicing).

On point ii): the iterative nature of the calculations and their very rigidly organized structure allows one to take full advantage of the vectorized features of a computer. Thus, on a Cray, for instance, sustained rates of tens of Megaflops can be achieved.

In conclusion a computing center which would serve the interest of high energy theorists should, of course, be endowed with one of the most powerful mainframes available, both in computational speed and in memory size.

Hadron Physics on a Lattice - An Example

The hadron spectrum shown in Figure 11 was calculated on a 6^2x12x18 lattice (where 18 is the time direction). The statistical ensemble used consisted of 20 Monte Carlo gauge configurations. The same ensemble was used to calculate a number of decay widths ($\rho\to\pi\pi$, $\Delta\to N\pi$, $\Sigma^*\to\Sigma\pi$, $\Sigma^*\to\Lambda\pi$, $\Xi^*\to\Xi\pi$, $K^*\to k\pi$). (Decay direction was made twice as long (12 sites) so that two hadrons could sit side by side, roughly speaking). Total CPU time for the calculation is about 5000-6000 VAX hours. Although this is a satisfactory start computing times between 10^3 to 10^5 larger are appropriate for this class of problem.

Comments:

1. As seen in the Figure 12, the lattice is both too small and too coarse.

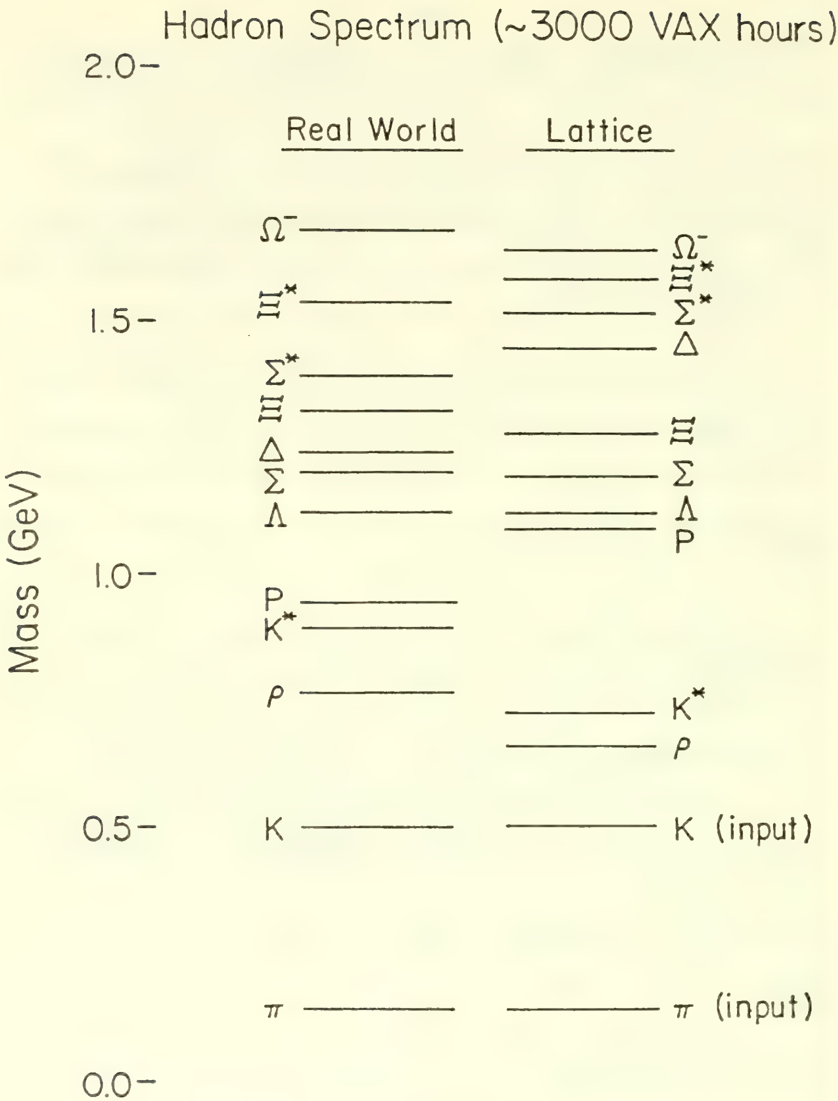

Hadron Spectrum (~3000 VAX hours)
Mass (GeV)
2.0
1.5
1.0
0.5
0.0
Real World
Lattice
Ω⁻
Ξ*
Σ*
Ξ
Δ
Σ
Λ
P
K*
ρ
K
π
Ω⁻
Ξ*
Σ*
Δ
Ξ
Σ
Λ
P
K*
ρ
K (input)
π (input)

Fig. 11

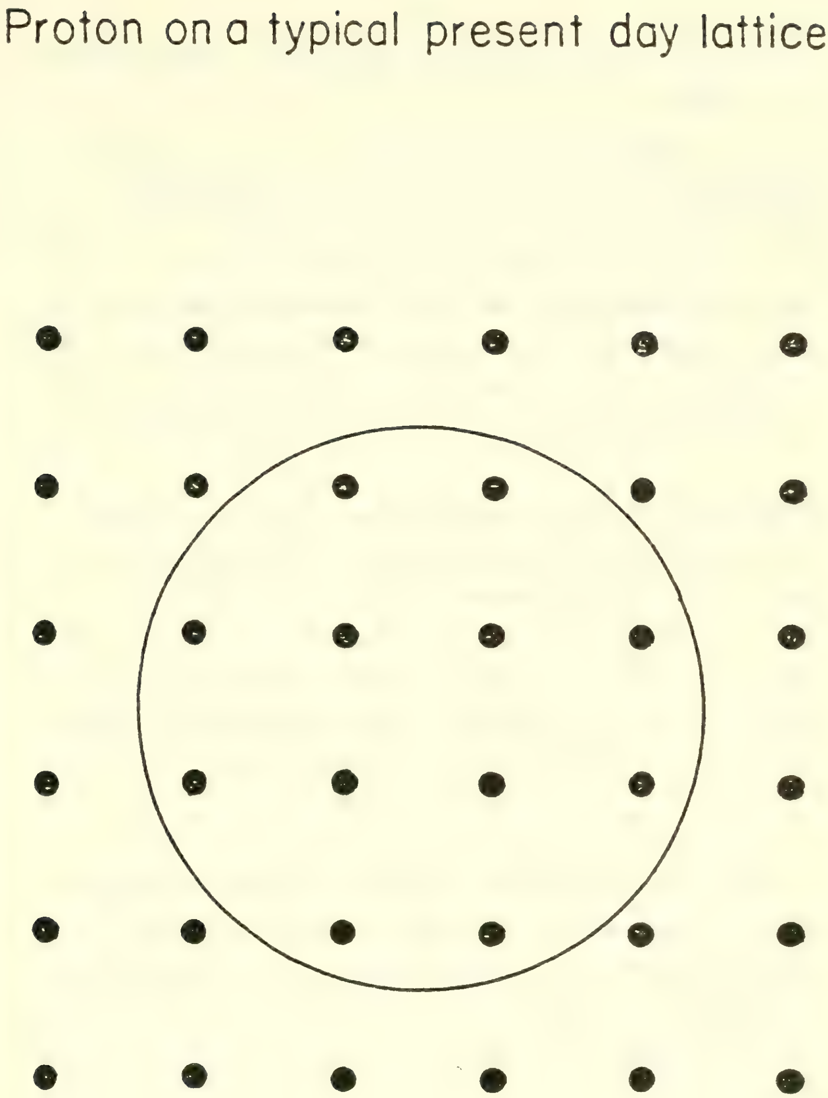


Figure 12

2. Closed fermion loop effects have not been included. (These are guessed to be ~20% effects.)

3. The lattice gauge hadron spectrum in Figure 11 is qualitatively encouraging, but still fairly crude on a quantitative level.

4. The above mentioned decays have been "observed" and measured on the lattice, but the number are even cruder than the hadron mass values.

5. Statistics are poor: 20 "events" $\rightarrow$ accuracy $\approx \frac{1}{\sqrt{20}} \approx$ 20-30%.

6. In principle, QCD is an <u>extremely accurate</u> theory. All the numbers it predicts should agree with the real hadronic world up to corrections of order $\alpha = 1/137 \rightarrow \lesssim$ 1% accuracy!!

(This is an extremely important point. QCD is not another arcane phenomenological model like Regge cuts, multiperipheral models, etc. It is fundamental in the same sense as QED. It's here to stay).

Nuclear theorists have made extensive use of computers for many years. In fact, many areas of contemporary theoretical nuclear physics depend critically upon large-scale numerical calculations. We include here a brief discussion of the relevant physics in areas where there is forefront activity and very active current interest.

Heavy-Ion Physics

In recent years, with the availability of new facilities, the experimental program has shifted from light-ion studies to heavy-ion physics. The computer demands of this field are orders of magnitude greater than those of light-ion physics. Collisions between complex nuclei confront the theorist with the problem of understanding nuclei under the unusual conditions of high excitation energy, angular momentum, deformation, isospin asymmetry and density. In these circumstances, a normal pattern of behavior is difficult to define and detailed realistic simulations are necessary.

At low energies (few MeV per nucleon), static simulations are important. Here one studies the potential energy and intertia of a nucleus as a function of its shape, temperature and argular momentum. Models based on macroscopic (liquid drop), microscopic (constrained Hartree-Fock) and macroscopic-microscopic (Strutinsky method) are typically used. To date, computational limitations (capacity and capability) have allowed exploration of only a small portion of the full parameter space involved.

The dominant microscopic quantum-mechanical theory for heavy-ion reactions is the time-dependent Hartree-Fock (TDHF) approach. Such calculations involve the evolution of single-particle wavefunctions for each nucleon of the colliding system and hence require large memory (several million words) and high speed. They are particulary suited to vector computation. A typical calculation for an intermediate mass system requires approximately 1.5 hours CRAY-1 time per impact parameter. Study of larger systems could easily consume 5 hours of time per week. There are various

ideas for extending the TDHF picture (functional integral methods, two-body collisions) but their realistic implementation is impossible with present computer capabilities. The reduction of the TDHF picture to a macroscopic nuclear fluid dynamics can only be achieved by a detailed comparison of microscoic quantum calculations with classical Liouville and fluid dynamical results in realistic situations.

The most recent Nuclear Science Advisory Committee Long Range Plan (1983) has given high priority for the construction of a Super-conducting Super Collider (SSC) as the premier new future facility. Such energetic heavy-ion collisions have some hope of seeing experimentally the existence of a gluon plasma. These high energies will require the development of new computational methods. They may involve nuclear cascade calculations, classical equations-of-motion methods, and nuclear hydrodynamics models. The goal would include an investigation of the extent of equilibration, the influence of the conditions for the existence of exotic phenomena. Since some of these calculations involve tracking the spatial evolution of each nucleon (or bit of matter) in the system; including mesonic and/or quark degrees-of-freedom very large amounts of computer time will be required. Relativistic formulations will exclude the usual simplifying approximation in the solution of the Schroedinger equation.

Nuclear Fission

The description of a fissioning nucleus involves quantum mechanical tunneling and extremely large amplitude collective motions. The recent development of a time-dependent mean field theory for spontaneous fission

provides a practical calculational framework for understanding the spontaneous decay of a many-fermion quantum system. As a generalization of the standard Hartree-Fock approximation each single-paticle wave function satisfies the time-dependent Schroedinger equation, with periodic boundary conditions in the time-dependent potential generated by all the other particles. The problem is reduced to the solution of a set of self-consistent integro-differential equations depending on three spatial variables and the time. The scope of such a calculation is substantial, requiring the iterative solution of partial differential equations for up to several hundred wave functions on a mesh in four space-time variables. Such calculations which might start with light systems such as Be^8, will ultimately want to be studied in U as well.

Many-body Problems and Field Theories

Stochastic techniques for the exact numerical evaluation of path integrals enable the exact solution of a variety of many-body problems and field theories of current interest. In the field of particle physics, this method opens up possibilities for exact solution of lattice gauge theory, which would make possible for the first time a quantum chromodynamic prediction for the structure and spectra of mesons and hadrons. In nuclear physics, these methods open the possibility of generating exact solutions to non-relativistic many fermion systems interacting via arbitrarily complicated two-body and many-body potentials, relativistic meson-nuclear field theories, and models formulated directly in terms of quark degrees of freedom.

The development of these techniques has been started in light nuclear systems in a truncated space. Ground state properties and spontaneous fission have been calculated in cases with two-body potential interactions with strongly repulsive cores. A nuclear theory containing vector and scalar mesons has been solved. Observable effects of quark substructure are being studied in the limitation of a one-dimensional confining model formulated directly in terms of quark-quark interactions. Initial efforts, utilizing stochastic evolution, have started for lattice gauge theories. New stochastic techniques are also being developed by particle theorists for more efficient calculations of quantum chromodynamics. As these developments progress and further insights and results are forthcoming, it is clear that future calculations will require larger computer resources than are currently available.

Another example of recent important efforts is the particle theory work at BNL in the numerical simulation in lattice gauge theory. These calculations represent an attempt to directly solve interacting field theories in four space-time dimensions. Some early results provide evidence that the gauge theory of quarks and gluons can simultaneously give rise to the phenomena of confinement.

This revolution in particle theory is made possible only because of the availability of modern computer facilities. Since the simulations are in four dimensions, the number of variables increases rapidly with the lattice size. In addition, the Monte Carlo techniques give statistical errors which decrease slowly with computer time. The results which are most certain involve the quantum dynamics of only the gluon fields and treat the quarks only as static

sources. Including the full relativistic quantum dynamics of the quarks is an active area of current research. A further increase in order of magnitude in computing power would allow calculations with the quantized quark fields. A large capacity vector machine is very important for attacking such problems on a realistic time scale.

The use of computers in particle theory is relatively new. This is changing rapidly and it is clear that the importance of numerical work in theory will grow quickly given adequate resources and support. It is likely that the computer will become as important a tool for the particle theorist as it is for the experimentalist.

4.2 BASIC ENERGY SCIENCE

Material Sciences

The development and proliferation of investigations of diverse material systems and phenomena via computer simulation and modeling is a rich field of scientific endeavour anchored in the physical sciences (with cross-fertilization links to advances in applied mathematics and computer science), made possible singularly by the advent of high-powered computers. Computer simulations provide information about phenomena and processes in material systems with refined microscopic spatial and temporal resolution and enable investigations of the dynamical evolution of complex systems under extreme conditions where data from experiments or other methods of investigation is not attainable. In addition such studies provide benchmarks

for critical testing and refinement of theoretical concepts and aid in the interpretation of experimental observations.

Current simulation methods involve the generation and analysis of phase-space trajectories of an interacting many-particle system either by the direct numerical integration of the equations of motion (molecular dynamics-MD, and reaction-trajectory-TJ, methods) or via the sampling of phase-space configurations (Monte Carlo-MC). In either case the many- body nature of the systems under study and the statistical modes of analyses dictate the necessity for extended computer time and storage capabilities.

The wide range of materials system investigated by computer simulations include: the equilibrim and non-equilibrium structure and dynamics of materials at different states of aggregation (solids and liquids) and the kinetics and dynamics of phase transformations; properties of metastable systems (supercooled liquids, quenched liquids, gasses); homo and multicomponent materials; ordered versus disordered (amorphous) solids; surfaces; interfaces and inter-phase interfaces, i.e, solid-solid (superlattices andcoherent structures), solid-liquid (epitaxial crystal growth and homogeneous nucleation), solid-gas (molecular beam epitaxy, heterogeneous catalysis).

Simulation studies on these systems allow investigation of structural and dynamical characteristics, kinetics and dynamics of phase-transformations, transport and non-linear phenomena (heat, matter, electrical), diffusion processes and reaction dynamics. Furthermore modifications of the intrinsic properties of condensed matter systems and phenomena (such as fracture, solid

transformations, plastic flow), due to external fields (mechanical stress, heat gradient etc.) can be investigated. In addition to an improved understanding of existing material systems, simulation studies could serve as the impetus for exploration of methods of preparation and growth of novel materials.

Underlying simulation studies of extended condensed matter systems is the notion that the properties of the "calculational sample" on which the simulation is carried out, extended via the commonly used periodic boundary conditions, are a faithful representation of the nature of the macroscopic system. Among the factors which dictate the size of the "calculational sample" are the ranges of interparticle interaction potentials and fluctuation wavelengths. Thus for example the MD simulation of the structural and dynamical properties of a solid simple metal (e.g., Al) requires a system containing ~2000 particles; the simulation of binary liquid metals and supercooled liquids require an even larger sample due to concentration fluctuations. Simulations of stressed crystals, fracture and plastic flow, shock wave propagation, the dynamics of melting and hydrodynamical phenomena would require systems where the number of particles would be 5000-10,000. It should be noted that in the presence of long range and realistic multibody forces the computing time grows as a (low) power of the number of particles. Such requirements necessitate memory capacity beyond CRAY-1S capability and large increases in computational speed.

A critical input in materials simulations is the interparticle interaction potential. A faithful simulation requires the calculation of such potentials via pseudo-potential methods which, for metalic systems, depends

upon the thermodynamic state variables (density, temperature, pressure). Simulations of nonequilibrium phenomena (such as solidification, quenching etc.) in which the state variables themselves evolve in time require a self-consistent adjustment of the interaction potentials along with the dynamical evolution of the system.

The coupled complexities of size and interaction potential calculations make such simulations prohibitive (and most likely impossible) on the current facility. Furthermore, the magnitude of such simulations dictate time allocations much beyond those currently available (for example, 80 minutes of CRAY-1 time allow the generation of ~5000 integration time steps for a system containing 1500 particles interacting via simple truncated Lennard-Jones potentials, with a fully optimized code. Note that this is the least demanding model from a computational point of view. A typical study of the solidification of such a system requires 50,000 integration time steps. It should be emphasized that the above considerations are dictated by the nature of the physical systems and phenomena and cannot be compromised by approximate treatments which will prejudice and distort the simulation results. Thus progress in this field cannot be accommodated without the substantial access to enahnced Class VI or Class VII computing facilities.

AB Initio Calculations of Materials Properties

Introduction

In the past few years, it has become possible to compute with good accuracy the detailed properties of materials and their surfaces without resorting to empirical means. Among the quantities obtainable from these calculations are crystal structures, static structural properties, vibrational properties, electron-phonon and phonon-phonon interactions, solid-solid phase transformations, surface electronic and geometric structures, and other static and dynamical properties. The general method is shown to be equally applicable to semiconductors, insulators, simple metals, and transition metals. With basically the atomic number and atomic mass of the constituent atoms as inputs, many of the above properties have been calculated to within a few percent of experiment. This development is made possible because of recent breakthroughs in the theory of total energy calculations and the availability of supercomputers. It has opened up many exciting possibilities for the study of condensed matter since one is in a position to predict properties of systems which were formerly inaccessible to theory or experiment.

Below, as examples, two applications of the theory are described. Both studies were made possible by the availability of CRAY time on the NMFECC. One involves the prediction of the geometry of the diamond (111) surface. The other is a calculation of the vibrational properties of bulk diamond and the phonon-phonon interactions in this material.

Theoretical Method

The general approach involves reducing the many-body problem to that of a set of self-consistent field equations for the electrons. For ground-state properties, this can be achieved in principle using the density functional formalism. The approach is ab initio in that the only inputs to the calculation are information about the constituent atoms and a subset of possible structural topologies among which a minimal energy structure is derived.

The solid is considered to be composed of rigid ion cores and the valence electrons which are itinerant. The electron wavefunctions are obtained by solving the wave equation using a basis set expansion. Interactions between the core and the valence electrons are evaluated using a pseudopotential which is constructed from a knowledge of the atomic number. The many-body electron-electron interactions are treated by the local density approximation in the density functional formalism. The core-core and electron-electron interactions are added to the core-electron interactions to give the total energy of the solid for a given structure. Once the lowest energy structure is found, the other solid state properties can be computed.

The method in principle is applicable to any system. However current applications have been restricted to systems with a limited number of atoms per unit cell. The reason is that the calculations involve setting up and diagonalizing large matrices for the solution of a secular equation. The computer memory required in a calculation is generally proportional to N^2 and the time required to N^3 where N is the number of atoms in the cell.

Diamond (111) Surface

The determination of surface structure, i.e. the arrangement of atoms on a surface, is a major unsolved problem in surface science. There is yet to be a definitive experimental probe for giving detailed surface geometry. For the diamond (111) surface, as for many other surfaces, there are many models proposed to explain the experimental data. Ab initio total energy calculations should be able to distinguish among these models and suggest a possible low energy structure.

In the calculation the structure is determined by minimizing the total energy with respect to the atomic positions of the first several layers from the surface. In addition to information on surface atomic rearrangements (i.e., relaxation with no change in surface symmetry or reconstruction with a change in surface symmetry), many other physical and chemical properties are obtained. Among these are surface electron wavefunctions and energies, work functions, surface energies, chemisorption geometries, bonding energies and charge distributions, and vibrational and optical spectra of the surfaces.

Table 4 summarizes some of the structural properties of bulk diamond calculated using a linear combination of atomic-like orbitals (basis set) for the electron wavefunctions. These results illustrate the accuracy of the method and serve as calibrations for the surface study. The computed lattice constant, cohesive energies, bulk modulus and its pressure derivative, and phonon frequencies are all in excellent agreement with experimental values.

TABLE 4. Ground State Properties of Diamond

Ground State	Experiment	Theory
Cohesive energy (in eV)	7.37	7.84
Lattice constant (in A)	3.567	3.560
Bulk Modulus (in Mbar)	4.42	4.37
Pressure derivative of bulk modulus	4	3.54
Zone-center optic phonon frequency (in cm^{-1})	1332	1344

The method was applied to study a variety of proposed reconstructions on the diamond (111) surface. This surface is of interest as the insulating limit for the group IV (111) surfaces which show a variety of surface reconstructions. Recent experiments indicate a hydrogen-terminated 1x1 surface at room temperature, but the 2x1/2x2 surface seen by Low Energy Electron Diffraction (LEED) for surfaces cleaned by annealing to above 1000 C is apparently H-free. Although there are many models proposed for this surface, the correct structure remains undetermined. In this study, energy minimization was carried out for all the topologically distinct models in the literature. These include the ideal relaxed model, the Haneman buckling model, the Pandey π-bonded chain model, the Chadi molecule model, and the Seiwatz single chain model. Of these only the Pandey π-bonded chain model

(which has chains of dangling bonds on the surface) has a lower energy than that of the relaxed 1x1 surface (Figure 13). A minimum-energy structure is determined for this model after extensive consideration of atomic position relaxations for all atoms in the first four layers. The other models are found to be implausible on the basis of their total energies and surface state dispersions as compared to photoemission results. The fully relaxed π-bonded chain model is a likely candidate in terms of energetic considerations and is consistent with all experiments reported to date. Also, contrary to previous belief, no dimerization of the surface chain is found to be favorable.

Phonon Frequencies and Anharmonic Coupling Constants in Diamond

The advent of total energy calculations makes possible an ab initio determination of phonon frequencies using a frozen-phonon calculation of the type shown in Figure 14. Here the potential energy vs. displacement for a zone center optical phonon polarized in the (111) direction is calculated. Each point (black dot) on the curve represents a calculation for one structure, namely a crystal with the phonon of amplitude Δu frozen in. The second derivative of this curve at the minimum gives the harmonic elastic constant and hence the phonon frequency.

This method has been extended to obtain, for the first time, an ab initio determination of higher-order anharmonic elastic coupling constants as well. If the curve in Figure 14 is expanded as a Taylor series about the minimum, then the third and fourth derivatives give the third and fourth order anharmonic coupling constants for this polarization. In order to extract these higher order derivatives reliably, it is necessary to repeat the

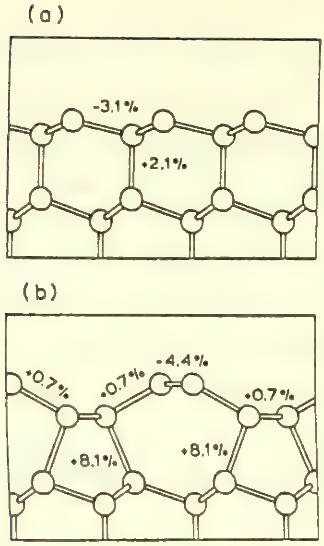


Fig. 13. Illustration of bond length changes (with respect to bulk) which occur upon relaxation of (a) 1×1 and (b) 2×1 Pandey chain models for the diamond (111) surface.

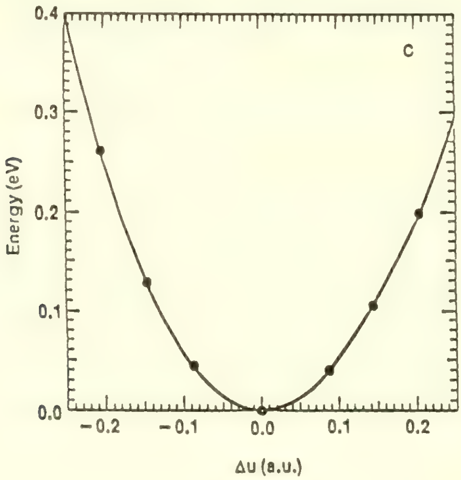
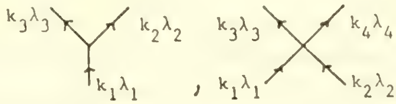
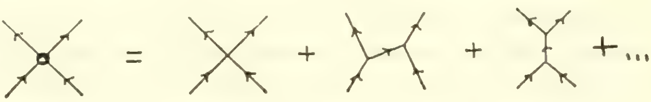


Fig. 14. Frozen-phonon energy vs. bond displacement.

calculations with greater accuracy and at more values of Δu. By looking at several polarizations, one can completely characterize the zone center optic coupling constants through fourth order. That is, the amplitudes for the three- and four-phonon processes



for $k \approx 0$ have been calculated from first principles. By making use of straightforward perturbation theory, one can even then calculate the renormalized 4-phonon vertex



in which exchange of an optic phonon is allowed. Somewhat less complete calculations for k-vectors away from zone-center, using a supercell technique, have also been computed.

A surprising result emerged from the calculation. It had been assumed that the fourth-order coupling was positive; in fact, it had been proposed as long ago as 1969 that a strong enough positive coupling could give rise to a two-phonon bound state that might explain an anomalous peak in the two-phonon Raman spectrum of diamond. Unfortunately, no reliable experimental or theoretical information has been available on the underlying coupling constants during the last 15 years, and the bound-phonon model has remained controversial. This calculation shows the renormalized coupling to be

negative, in which case no bound state could form. It is therefore believed an alternative explanation for the Raman anomaly is probably the correct one.

Computing Requirements

The work described here, namely the calculation of surface structures and of phonon coupling constants for diamond, requires substantial computer resources. The most demanding parts of the calculation are the solution of the Schroedinger secular equation, which requires the diagonalizataion of ~300x300 complex matrices, and the determination of the charge density, for which ~40x40x40 3-d Fourier transforms must be carried out. These operations must be repeated frequently as one loops over bands, k-vectors, self-consistent iterations, and test geometries. Both parts of the program lend themselves to vector machines, and have been optimized to make efficient use of the CRAY pipeline architecture.

The calculations required approximately 30 hours of CRAY computer time in total. A typical calculation for a given test geometry for the surface case required 25 minutes of CRAY time. Under typical turn-around conditions, such a job might take a few hours to complete. The same job could in principle have been run on a smaller machine, e.g. on an IBM 370/168; it would have required ~5 hours of CPU time, or perhaps a day under similar turn-around conditions. The same job would take roughly 2 days of CPU time alone on a VAX 11/780. The availability of reasonable turn-around for such CPU-intensive jobs is a strong motivation for the use of Class VI computers.

In addition to CPU, precision and memory are factors. Essentially all the calculations are done using full (~64 bit) precision. The memory requirements currently extend up to ~4 Mbyte, which is too large for some machines (e.g. the CDC 7600).

In summary, the combination of CPU-intensive, vectorizable, high-precision, and large memory requirements makes a Class VI computer the obvious choice for work of this kind. The above studies are only two examples among the many applications. Future applications are expected to be numerous and exciting. Current studies are restricted to rather simple systems because of computer capacity and capability limitations. For example, transition metal surfaces with complex adsorbates have not been attempted (even with access to CRAY 1 computers) because of the extensive demand of computer time and memory for those calculations. With the future availability of Class VII or better supercomputers this general approach should have the power to predict the existence of new materials, novel properties and phenomena, and to extend the study of materials to physical situations (such as those at complex surfaces and interfaces) which are difficult or impossible to examine experimentally.

AB Initio Alloy Theory

Introduction

The development of new and useful alloy systems and the elucidation of their properties is an area of enormous technological importance. A marginal increase in the allowed operating temperature of alloys used in commercial

power plants, improvements in strength to weight ratio of allows used in the aero-space industry, improvements in the corrosion resistance of alloys used in chemical process plants would all produce massive savings both in terms of the cost of the final product and in terms of the demand for raw materials.

Traditionally, the search for new alloy systems has been conducted largely on a trial and error basis, guided by the skill and intuition of the metallurgist, large volumes of experimental data, the principles of 19th century thermodynamics, and ad hoc semiphenomenological models. Recently this situation has begun to change as a result of developments that are making it possible to calculate the physical and metallurgical properties of alloys from first principles. It is thus becoming possible to understand the underlying physical mechanisms that control the formation of alloys and determine their properties, and, thus, for theory to offer guidance in alloy design.

Within the content of alloy theory one is interested in the whole phase diagram of the alloy system, the pure metal, and points where the possible ordered phases can form, and the possible solid solutions. For pure metals and ordered alloys the role of the modern supercomputers has been in extending the capability (in terms of the number and type of atoms in a unit cell) of the system that can be treated. In the theory of random solid solution, the use of supercomputers has played an integral role in the development of the theory even at the level of its application to simple binary alloy systems. In the following an explanation of why this is so and examples of the kinds of important problems in alloy theory that have been treated are presented. In these examples the availability of a supercomputer was vital to the completion of the project.

Electronic Structure Calculations for Random Alloys

The fact that random alloys do not possess translational invariance does not affect (at least in principle) the applicability of the density functional methods, it does however mean that the band theory methods developed to solve the Schrödinger equation for periodic systems are no longer useful. During the last few years, however, methods have been developed for calculating the electronic properties of random alloys which, though computationally difficult, allow them to be treated on roughly the same footing as pure metals and ordered alloys.

The central problem in dealing with random alloys is that the effective one electron crystal potential function entering the Schrödinger equation is random. The study of the eigenvalue spectra of random Hamiltonians has a long history which has resulted in the development of approximations of known standing which can be used to treat specific problems. One of the major developments is the so-called coherent potential approximation (CPA). The CPA allows the treatment of the effects of disorder at a level that makes it worthwhile to attempt to use it as the basis of a first principles theory of random alloys.

The modern first principles theory of the electronic structure is based on a method called the Korringa-Kohn-Rostoker-CPA (KKR-CPA). In this method the effective one-electron crystal potential is treated at the same level as that in ab-initio band theories of ordered metals (the so called muffin-tin model) indeed the KKR-CPA method becomes for an ordered solid the KKR-band

theory method, one of the standard band theory methods used for treating metallic systems.

It is a measure of the added complexity involved in solving the Schrödinger equation with random potentials that the first fully self-consistent calculation of the electronic structure of a random alloy using the KKR-CPA method was published in 1982 and was performed on a CRAY-1S supercomputer.

Short Range-Order (SRO) and Clustering In Alloys

The occurrence of short range order (the preference of one alloying species for neighbors of a different species) and clustering (the preference for like neighbors) in obstensively random alloys can substantially affect material properties. It is, therefore, of considerable importance to be able to understand the physical mechanisms that drive ordering processes in alloy systems. With the development of first principles methods of calculating the electronic structure of random alloys it is becoming possible to develop models that are capable of predicting the occurrence of SRO and clustering.

A mathematical description of this class of problems requires a proper treatment of both the electronic structure of the alloy as well as the statistical mechanics of alloy configurations. Within the context of first principles methods such models are computationally demanding. The following are examples of two different, though complementary, methods that are currently being developed for treating SRO and clustering in alloys. Both

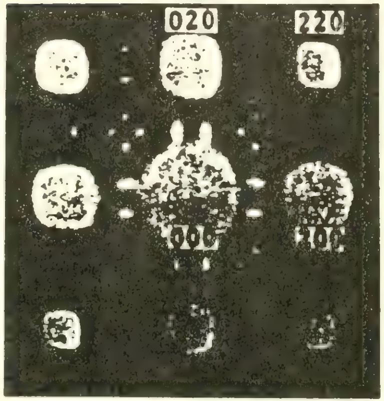
problems have relied heavily on the availability of a CRAY-1S supercomputer both for the calculation of the electronic properties of the homogeneous random alloy method and for the particular treatment of SRO.

Concentration Functional Method

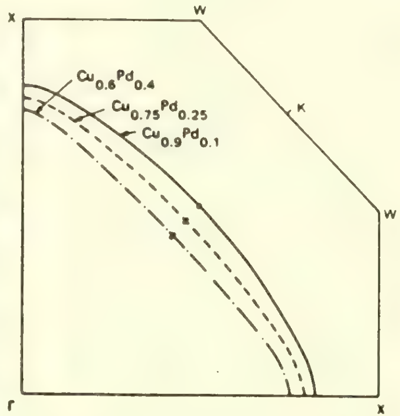
This approach is based on a classical analogue of the density functional approach. It has been used to explain the concentration variation of the wave vector, $\vec{k}_0$, of concentration waves in Cu_cPd_{1-c} alloys. Concentration waves may be thought of as long-lived periodic fluctuations of the site concentration that modulates the disordered state. Concentration waves are observed directly as diffuse scattering peaks in x-ray, neutron-, and electron-scattering experiments. (See Figure 15) For the Cu_cPd_{1-c} alloy system as well as many similar systems the separation between the diffuse scattering peaks varies rapidly with concentration.

It was found the concentration-dependent diffuse scattering peaks that are observed in Cu_cPd_{1-c} alloys arise for values $\vec{k}_0$ that connect well-defined parallel flat pieces of the alloy's Fermi surface. Thus in this system the SRO is electronically driven. In Figure 15 are shown the Fermi surfaces based on the results of KKR-CPA calculations of the electronic structure, of three Cu_cPd_{1-c} alloys in the rXWK-plane of reciprocal space. Also shown in this figure is a comparison of the measured and calculated concentration variation of the separation, m, between the diffuse scattering spots.

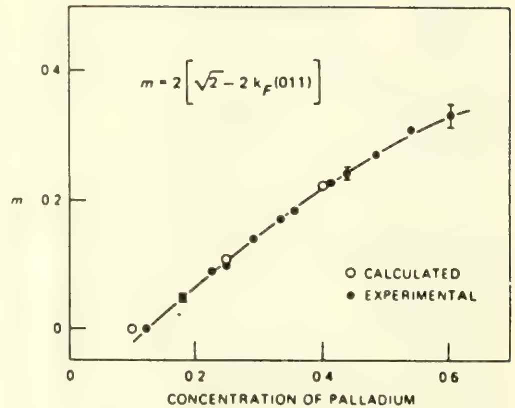
A NEW THEORY OF ORDER-DISORDER PHENOMENA IS USED TO EXPLAIN THE OCCURRENCE OF CONCENTRATION WAVES IN SUBSTITUTIONAL Cu_cPd_{1-c} ALLOYS



DIFFUSE SCATTERING FROM CONCENTRATION WAVES IN A DISORDERED $Cu_{0.67}Pd_{0.33}$ ALLOY



CALCULATED FERMI SURFACES IN Cu_cPd_{1-c} RANDOM ALLOYS



CONCENTRATION DEPENDENCE OF DIFFUSE SCATTERING SPOT SEPARATION IN Cu_cPd_{1-c}

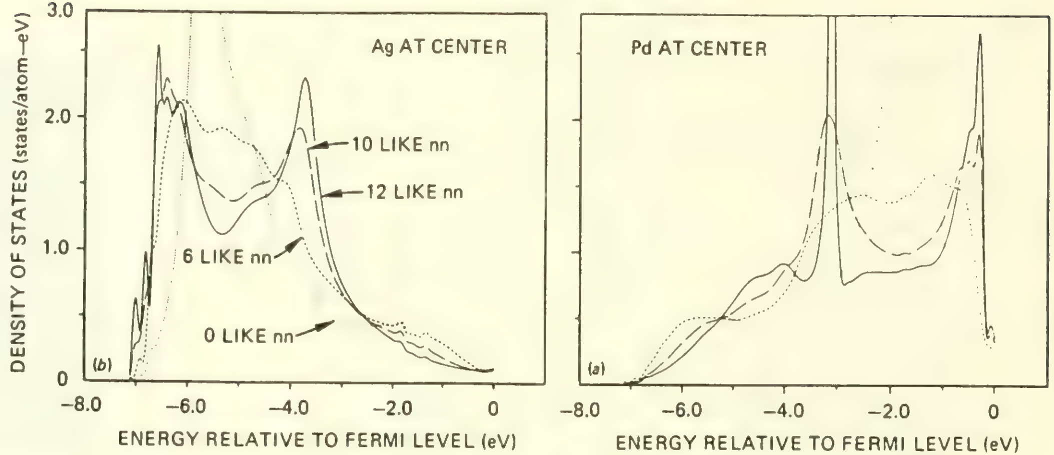
Figure 15

Cluster Energies

A second approach to the question of ordering in random alloys is to consider the energies associated with particular local arrangements of clusters of atoms within an otherwise homogeneous random alloy. The results of such calculations can then be used as input paramters for a proper treatment of the statistical mechanics.

The calculation of cluster energies first requires that the CPA theory be extended to take into account the effects of specific local configurations. In Figure 16 are shown the first such cluster densities of states that have been calculated within a first principle theory. In Figure 16 calculated densities of states for palladium and silver sites at the center of particular 13-atom clusters in an $Ag_{0.5}Pd_{0.5}$ random alloy are compared with those for the random alloy. Clearly, the cluster density of states can deviate substantially from those obtained for the homogeneous random alloy. In particular, the density of states $n(\varepsilon)$ for the central site of an all-palladium cluster (solid line) resembles that of pure palladium, but $n(\varepsilon)$ for a palladium atom surrounded by 12 silver atoms displays a single peak (dotted line) characteristic of an impurity atom embedded in a pure metal. On the basis of preliminary calculations of the cluster energies the results indicate that unlike near neighbors are energetically favored over like near neighbors; i.e., this alloy would tend toward SRO rather than clustering.

CLUSTER DENSITIES OF STATES IN $Ag_{0.5}$ $Pd_{0.5}$ FOR A CLUSTER CONSISTING OF A CENTRAL SITE PLUS 12 - NEAR - NEIGHBORS (nn)



- CENTRAL SITE DENSITY OF STATES VERY DEPENDENT ON nn OCCUPANCY

- 12 - LIKE nn GIVES $n(\epsilon)$ SIMILAR TO PURE METAL
  0 - LIKE nn GIVES IMPURITY LIKE $n(\epsilon)$
  6 - LIKE nn GIVES $n(\epsilon)$ SIMILAR TO CPA

Fig. 16

Chemical Vapor Deposition Modeling

It is becoming increasingly feasible to model chemically reacting flows in considerable detail. Many examples are found in combustion systems, especially those concerned with pollutant formation. Here, a materials science example is used to illustrate the need for increased large-scale scientific computing capability. Even the Class VI computers currently in use are sometimes inadequate in both speed and memory.

The chemical vapor deposition (CVD) of solid materials is an important method for producing solid films with high purity and uniformity. Important applications include the manufacture of microelectronic devices and silicon solar cells. Even though various forms of the process are used commercially, relatively little is known about the basic physical and chemical behavior of the process. As a result, considerable difficulty is being experienced in the reliability and the scaling-up of the manufacturing process, especially when new chemicals are introduced. Precise information is required on the effects of parameters such as carrier gas composition, flow rates, and temperatures on results such as deposition purity and uniformity.

Chemical vapor deposition is a process in which a gas, say silane (SiH_4), flows over a heated substrate. As heat is transported away from the substrate it heats the gas, causing it to decompose or pyrolyze. Some of the species that are formed by the pyrolysis, SiH_2 and Si_2H_2 for example, react with the heated surface, depositing solid silicon and releasing hydrogen back into the gas phase. Modeling the process requires solving conservation equations that

describe the chemical reaction and the convective and diffusive transport of mass, momentum, energy of 30 or more chemical species.

A mathematical model has been constructed that describes the process faithfully enough to include most of the features of interest in this process. It is formulated in terms of elliptic partial differential equations. Solution of systems of these equations is not feasible on a Cray 1 because it requires keeping in memory vast amounts of information. At each grid point in the entire spatial region, 30 to 900 items of information are required. To solve the elliptic equations it is necessary to keep on hand information at all grid points for the entire region. Since that cannot be done (32 megawords would be required), a "boundary layer" approximation is made. This technique yields a new set of partial differential equations to be solved that are parabolic instead of elliptic. With these equations it is possible to keep much less information on hand. The penalty is that phenomena in several subregions of interest are not modeled adequately. In particular, important effects along the leading edge of the heated graphite susceptor are ignored by the simplified model.

CVD modeling is further limited by Class VI computers in that many chemical species of great practical interest cannot be included. Presently there is much interest in chloro-silane species. From a modeling standpoint, adding chlorine to the system increases dramatically the number of species that have to be considered, and thus the size of the problems. Moreover, thermochemical properties must be determined for most of the new species.

Thermochemical properties (specific heats, formation enthalpies, etc.) of many important species involved in CVD processes are not known, and because they are short lived and present only in trace amounts, their properties are difficult to measure. Nevertheless, precise information on thermochemical properties is critical to the success of CVD models. Fortunately, it is now possible to compute the electronic structure of molecules, and thus determine their thermochemical properties. The electronic structure models have especially large computer memory requirements. Even with Class VI computers with two to four megawords of storage, the models spend most of their "wall clock" time moving information to and from disc. A typical computation that requires 1 hour of CPU time will often take over 100 hours to complete because of the swapping caused by insufficient memory. Increasing memory to 20 to 40 megawords would result in nearly two orders of magnitude decrease in turn around time for these important computations (Figure 17).

Molecular Structure and Properties By Monte Carlo

Although Monte Carlo methods have a long history in the physical sciences, until recently most applications have been to non-quantum mechanical situations. In the past several years, however, there has been a surge of interest in computational methods which are not only quantum-mechanical (e.g., calculating expectation values with a given trial wave function), but which truly simulate the underlying probability distribution of a many-body system. In principle, by this approach quantum-mechanical expectation values can be calculated exactly. Although there has been some experience with the

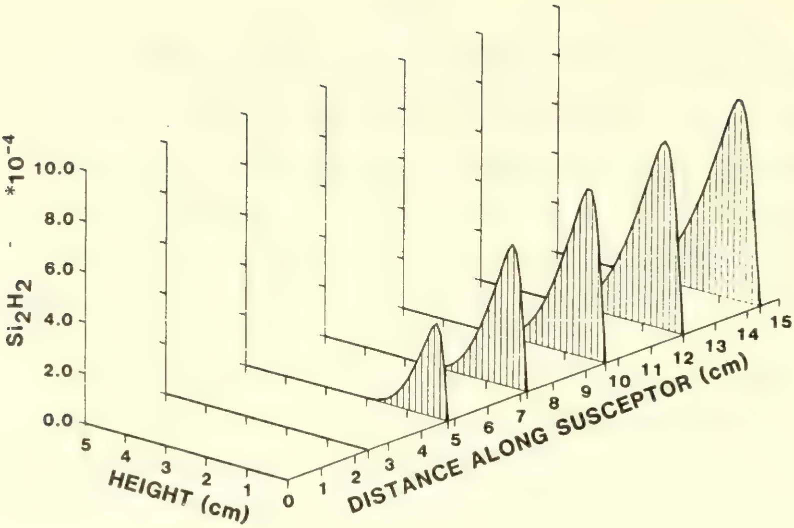


Figure 17. Gas phase Si2H2 profiles as a function of distance along the susceptor and height above the susceptor. The flow entering at the left is a mixture of SiH4 and He. The gas is heated as it flows along the susceptor, and species such as Si2H2 are formed by pyrolysis. These pyrolysis products react with the susceptor to deposit solid silicon, and release hydrogen into the gas phase.

many-Boson problem, the difficult but physically important many-Fermion problem is only now being "cracked".

These developments are currently being applied in molecular physics and quantum chemistry enabling highly accurate calculation of the correlation energy of many-electron systems. In addition, chemically more significant quantitites such as binding energies, barriers to chemical reaction, level splittings, and other properties are being obtained to within chemical accuracy. In particular, highly accurate correlation energies have been obtained for He, H_2, LiH, Li_2, N, CH_2, H_2O, and N_2; the classical barrier height for the H + H_2 exchange reaction has been computed to within 0.1 kcal/mole; and the singlet-triplet splitting in methylene, the binding energy of N_2, as well as several points on the potential energy surfaces of Li_2 and H_3 have been obtained.

One way to approach such quantum Monte Carlo simulation is to re-write the time-dependent, many-body Schrödinger equation in imaginary time. The resulting equation is a multi-dimensional diffusion-type equation, which can be shown to have a steady-state solution which is the ground state of the time-independent Schrödinger equation. Simulation of an appropriate random-walk process then allows one to obtain molecular properties by simply "measuring" them as the system "evolves". The system is in its steady-state when the distribution of walks is in equilibrium.

For most chemical applications, the Bose and Fermi ground states are well separated in energy. In this case it is useful to use the so-called "fixed-node" approximation for handling the identical particle antisymmetry

(Pauli exclusion principle). This is a highly accurate approximation in which the energy satisfies a variational principle in the nodes of the trial wave function. In Figures 18-20, are shown some properties of the nodes of H_3 obtained with a single-determinant trial wave function which is multiplied by a Jastrow correlation factor. Figure 21 shows the result of using different time-steps τ for propagatinng the random-walk forward. The approximate form for the Green's function results in a small bias which vanishes as $\tau \rightarrow 0$. In most situations the time-step bias is quite small (cf., Figure 21). Where it is significant, one may extrapolate to $\tau \rightarrow 0$, or use the fully exact Green's function Monte Carlo methods of Kalos, et al.

Monte Carlo methods are notoriously CPU intensive. Advances in theory have led to major speed-ups in the basic algorithm. Nevertheless, reduction of the statistical error into the necessary range is a costly process. This is because the statistical uncertainty σ decreases slowly, as $1/\sqrt{N}$ for N sample points, with N proportional to the CPU time. (In principle, it is possible to reduce the error faster by using information obtained during a run to improve the importance function being used. However, such algorithms are still under development.) The statistical precision required in chemical applications is far more severe than in most other applications in the physical sciences. Although an accuracy of 99.9% is generally deemed quite exceptional, in chemistry, where the quantities of interest are small <u>differences</u> of large numbers, the 0.1% uncertainty can easily mask the entire difference. (Methods for obtaining Monte Carlo estimates of energy differences directly are also possible. Work along these lines is currently being actively pursued.) Reduction of the uncertainty by an order of magnitude (e.g., to 0.01%) costs 100 times as much CPU time.

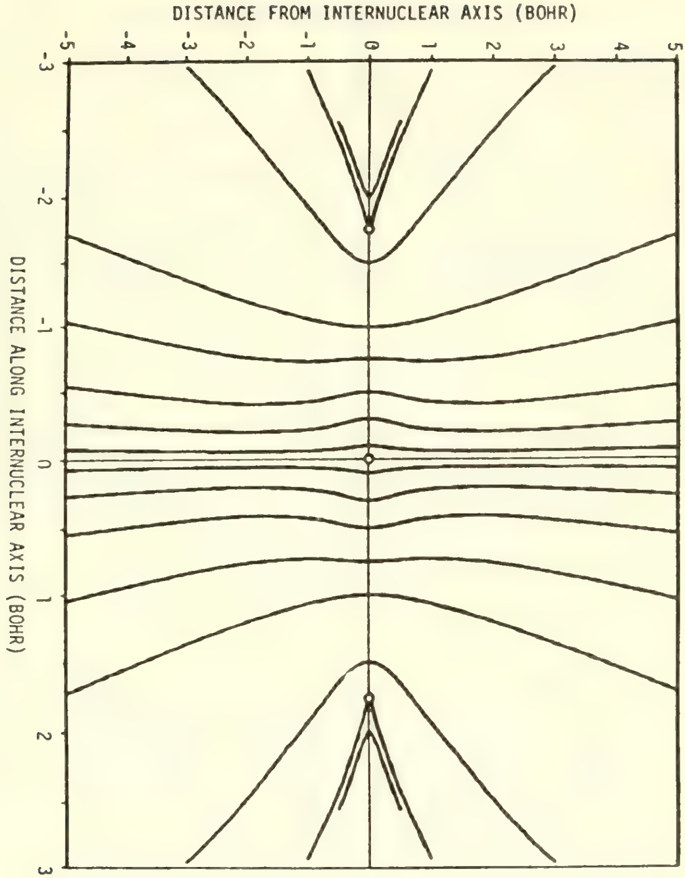


Figure 18. Nodes of a single-zeta quality importance function for the H_3 system at its saddle-point geometry. The nodes (zeros) of the importance function determine the ultimate accuracy attainable by Monte Carlo within the fixed-node approximation. Here the three circles along the x-axis are the three hydrogen nuclei. The curves are cross sections of a representative subset of surfaces along which the importance function vanishes when two like-spin electrons are both on one surface.

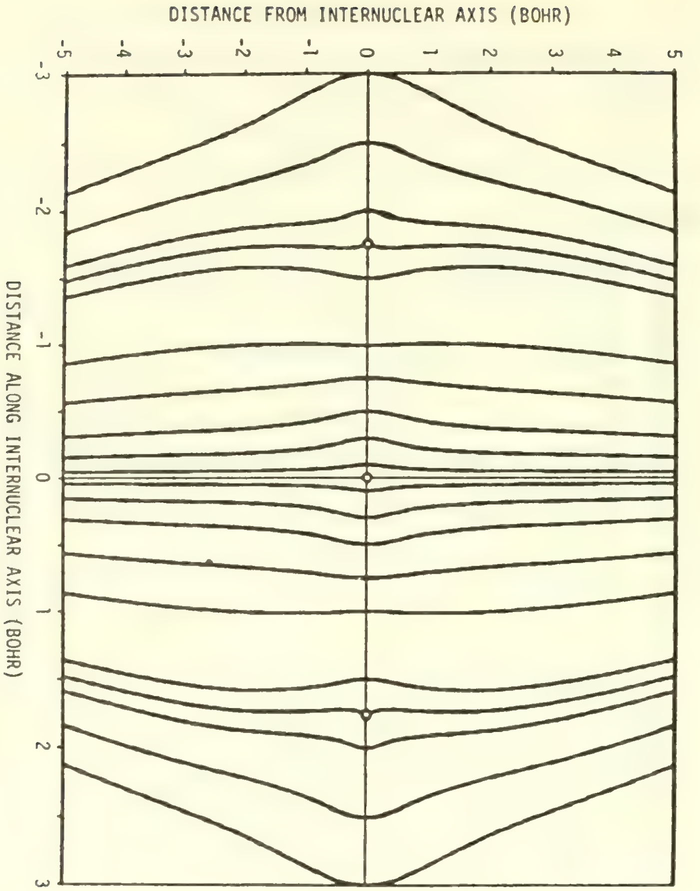


Figure 19. Nodes of a double-zeta quality importance function for H_3 at its saddle-point geometry. See Fig. 18 for further discussion. It is clear from these figures that single- and double-zeta basis sets result in significantly different nodal surfaces. Further basis-set enhancement, however, leads to no significant change in the nodes. In fact, importance functions approaching Hartree-Fock quality have nodal surfaces that are almost indistinguishable from those shown in this figure.

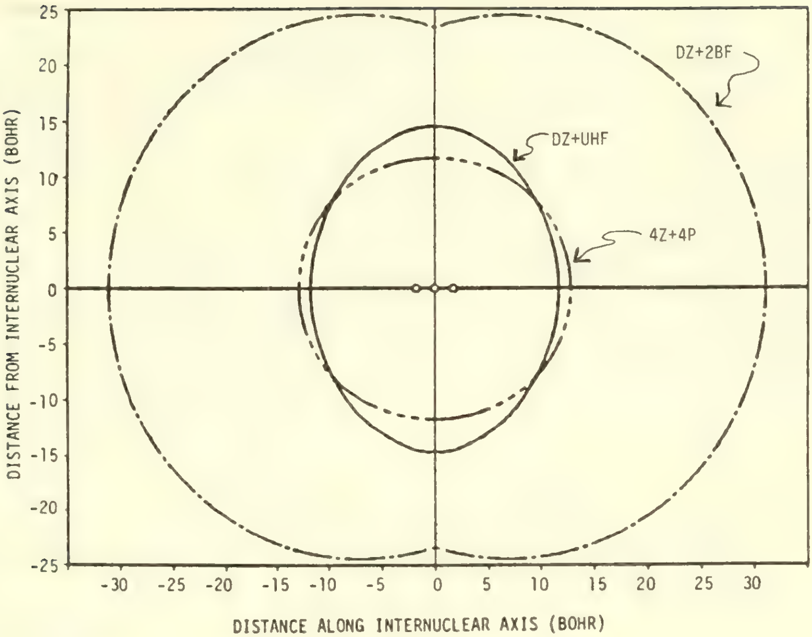


Figure 20. Nodes of the lowest-energy molecular orbital of H_3 at its saddle-point geometry for three different basis sets. In addition to the particle exchange nodes shown in Figs. 18 and 19, one additional nodal surface occurs in the importance function as a result of a node in the $1\sigma_g$ molecular orbital. The importance function vanishes when both like-spin electrons are on this surface <u>or</u> when the single unlike-spin electron is on this surface. As can be seen, the scale of this figure is much larger than that of Figs. 18 and 19. The three nuclei all appear very close to the origin. The outermost curve is for a basis set consisting of two basis functions on each nucleus (double zeta), denoted DZ, plus two functions centered between the nuclei (DZ+2BF); the curve labelled DZ+UHF is for a spin-unrestricted DZ importance function; 4Z+4P denotes a near Hartree-Fock limit basis at the level of four zetas plus four polarization functions.

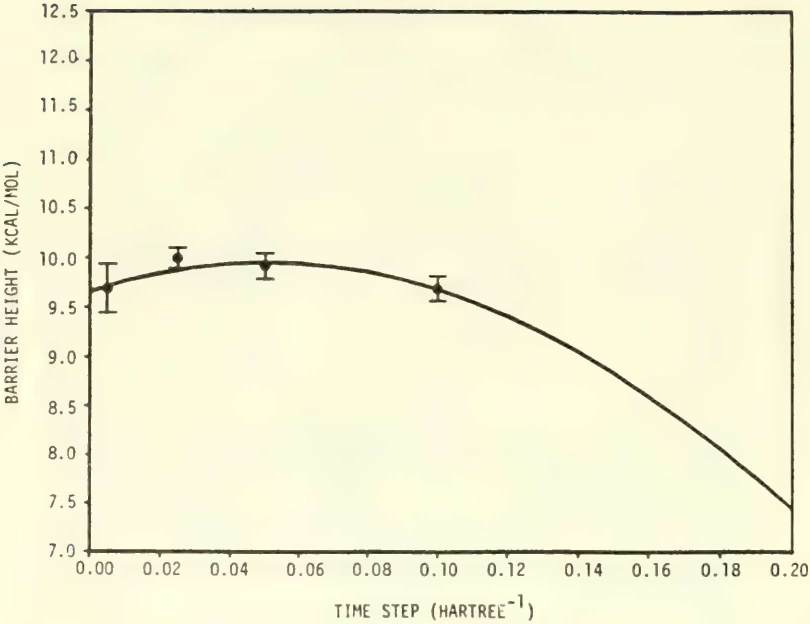


Figure 21. Extrapolation of the $H + H_2$ barrier height to zero time-step size. The short time approximation to the Greens function results in a small bias for finite iteration times. An unbiased estimate may be obtained by extrapolation. The four points and associated error bars are Monte Carlo calculations of H_3 at its saddle-point geometry. The curve is a polynomial least-squares fit to these points to order x^2. For H_3 it is clear that the final result is relatively insensitive to the time-step size.

The above considerations make Monte Carlo calculation of chemical properties an activity that demands much Class VI computer time. Furthermore, treatment of relatively large molecules will require Class VII computers and beyond, as the CPU time required scales roughly as n^2, where n is the number of electrons in the system. (This is significantly better than traditional quantum chemical methods which scale as the 4^{th} to 5^{th} power of the number of basis functions, which is roughly proportional to the number of electrons.) With sufficient Class VII computer time one would be able to tackle a number of important problems in areas such as combustion research, which have thus far been beyond computational reach.

Molecular Design Principles For Biomimetic Solar Energy Conversion Systems

Introduction

Studies aimed at the development of biomimetic solar energy conversion systems have increased significantly over the last decade. This has been due in large part to a concommitant increase in our knowledge of the details of the photosynthetic process in both plants and bacteria. However, while increased knowledge of the photosynthetic process has certainly aided the search for appropriate biomimetic systems, significant problems remain.

For example, in photosynthetic bacteria, conversion of light energy into chemical energy begins with excitation of a bacteriochlorophyll dimer, $BChl_2$, to an excited singlet state, followed by rapid electron transfer from $BChl_2$ to a bacteriopheophytin (BPheo, a demetalated BChl) acceptor,

$$BChl_2 \cdot BPheo \rightarrow BChl_2^{*} \cdot BPeo \rightarrow BChl_2^{+} \cdot BPeo^{-}$$

This scheme suggests that viable artificial solar energy conversion systems might be based on donor-acceptor complexes constructed from similar chlorophyll-like components, which are able to form relatively long-lived excited charge-transfer (CT) states. However, practical means for determining the precise nature of the components and their overall molecular architecture do not exist at present. And due to the large number of molecular possibilities and to the difficulty of actually constructing such complex systems, the need for a coherent set of design principles is paramount.

In the following, we shall provide an example of the way in which molecular quantum mechanical methods can play a significant role in the development of such molecular design principles. The example derives from a series of covalently-linked "face-to-face" porphyrin dimers originally synthesized by C.K. Chang (J. Heterocyclic. Chem. 14, 1285 (1977)), and is based on the prototypical model porphine-magnesium porphine dimer, P-MgP, depicted in Figure 22. Energy trapping via charge separation in this system may be thought to occur as a two-step process:

1. excitation of P-MgP to an excited (π,π^{*}) singlet state, followed by

2. internal conversion to S_1, an intramolecular radical-pair CT state of the type P^{-}-MgP^{+}. Ideally, this state should be relatively stable towards decay to the ground state, S_0.

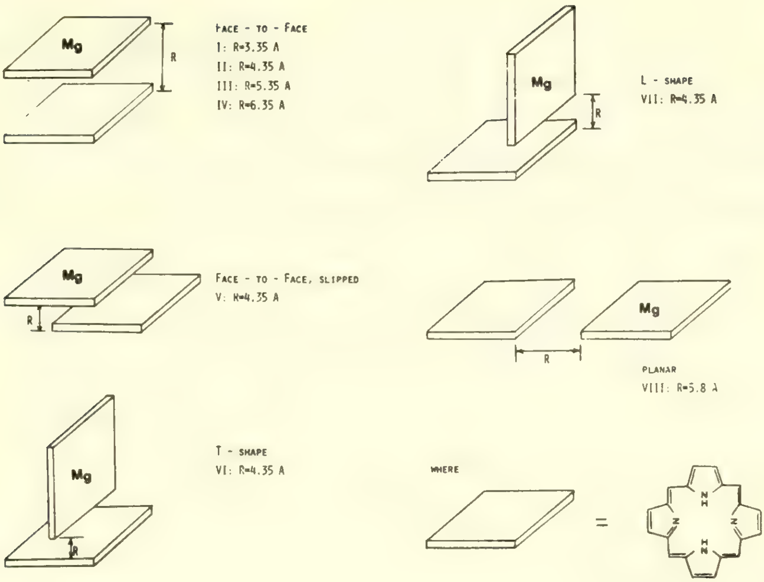


Figure 22. The development of biomimetic solar energy conversion systems rests heavily on the concept of photo-induced charge separation. In the present study, the asymmetric porphyrin dimer, P-MgP, was investigated due to the general similarity of its molecular components to those of the chlorophyll species found in the photoreaction-centers of plants and bacteria. As the exact geometric arrangement of the molecular components found in biological photoreactin-centers is unknown, a number of possible dimer geometries were studied in order to investigate the relationship of dimer structure to formation of low-lying CT states, and to determine whether a particular geometry was significantly better than the others for CT state formation. Data obtained in the current work strongly suggests that face-to-face, closely-spaced dimers appear to be the best candidates for the development of porphyrin-based biomimetic solar energy conversion systems.

Experimental studies by Netzel et al. (Chem. Phys. Letts. 67, 223 (1979); J. Phys. Chem. 86, 3754 (1982)) have characterized a number of the photodynamical properties of a closely related system under a variety of conditions, and the relevance of their experimental results to our calculations will be addressed in more detail shortly.

A key aspect of both the theoretical and experimental studies involves investigation of the relationship of the geometric and electronic structural features of the dimers to their photodynamical properties, and how these properties are affected by the nature of the surrounding microenvironment. The results thus obtained provide a basis for determining under what conditions the P-MgP system will likely act as an effective phototrap.

Theoretical Methodology

The work presented here is a theoretical study of the excited states of P-MgP performed with *ab initio* Hartree-Fock self-consistent field molecular orbital (SCF MO) and configuration interaction (CI) methods, using a basis set of floating spherical Gaussian orbitals (FSGO). The calculations were performed for the molecular geometries shown in Figure 22, with each complex treated as a supermolecule; the results include Franck-Condon transition energies, $S_n \leftarrow S_0$ oscillator strengths, and charge distributions of a number of the low-lying excited singlet and triplet states.

Computational Considerations

Ab initio electronic structure calculations on molecular systems the size of those studied in the current work present significant computational problems. The problems involve the need for high speed processing, high speed I/O, and large memory and disk file space. All calculations described herein were carried out on Class VI level CRAY 1S computers at United Telecom Information Systems in Kansas City, Missouri and at NMFECC in Livermore, California.

A complete calculation on one specific molecular geometry (c.f. Figure 22) requires several computational steps, some of which present potential computational bottlenecks. First, the SCF MO calculation on a single P-MgP dimer consists of the evaluation and storage of a large number of electron repulsion integrals, followed by iterative processing of the integral files to produce a set of converged molecular orbitals. The number of integrals, I, is related to the number of basis functions (*viz*. FSGO's), N, needed to describe the molecular electronic structure by the following equation:

$$I = N^4/8 + N^3/4 + 3N^2/8 + N/4$$

For P-MgP where N = 249, I = 484 398 375; however the majority of integrals are small in magnitude (greater than 95% of them are $< 10^{-9}$) and need not be evaluated explicitly. Nevertheless, a typical dimer may require evaluation of 15 - 40 million integrals. Table 5 below provides a representative sample of execution times and storage requirements for integral evaluations on a

CRAY 1S. It is seen that while the CRAY 1S is adequate for present calculations, additional resources and speed will be necessary for systems in which 80 - 100 million integrals must be computed. Significant computational bottlenecks can also arise when smaller numbers of integrals must be computed a large number of times, such as is required in geometry optimizations. Moreover, it should be noted that the SCF MO calculations, which due to their iterative nature place a substantial I/O burden on the system, require an additional amount of time equal to approximately 70% of the integral evaluation time.

TABLE 5

Number of Integrals (Millions)	Integral Storage (Mbytes)	Integral Comput. Time (Minutes)
15	180	11.5
25	300	19.2
40	480	30.7
60	720	46.1
80	960	61.5
100	1200	76.9

Calculation of excited-state wavefunctions by a CI procedure presents an additional computational bottleneck due to the integral transformation needed to transform electron repulsion integrals over basis functions to integrals over molecular orbitals. The time, t, is given by

$$t \sim M * I$$

where M is the number of molecular orbitals to be transformed, typically 30 - 50. Thus, this step is potentially more costly than the entire integral

evaluation. However, we have developed highly vectorizable code for the transformation which, when executed on a vector-processing machine such as a CRAY 1S, reduces t to a value considerably less than than required for integral evaluation. It must be emphasized that without the vector processing capability of a CRAY 1S or without other state-of-the-art computational enhancements (e.g. highly parallel processor architecture), execution of the integral transformation would be extremely time consuming and prohibitively expensive for studies such as those described here.

Isolated P-MgP Dimers

In calculations on isolated P-MgP for all but geometry I, the supermolecule molecular orbitals closely resemble those of the constituent monomers, with only modest mixing of electron density from the perturbing ring. Consequently, as shown in Figure 23, the lowest four computed excited states, relative to the ground singlet state S_0, appear as slightly perturbed monomeric (π,π^*), states, with excitation largely confined to a single ring. S_1 and S_4 resemble the two lowest excited states of porphine, labeled $Q_y(P)$ and $Q_x(P)$, respectively, while S_2 and S_3 correlate with the doubly-degenerate Q(MgP) state of magnesium porphine. Within the accuracy of the computational method, these states are not appreciably shifted from their monomeric transition energy values.

The lowest two computed CT states, $S_5(CT)$ and $S_6(CT)$ are shown in Figure 23 to be much higher above S_0 than the four lowest (π,π^*) states. $S_5(CT)$ and $S_6(CT)$ are characterized as singlet radical-pair states of the type P^--MgP^+, and correspond essentially to excitation of an electron from the

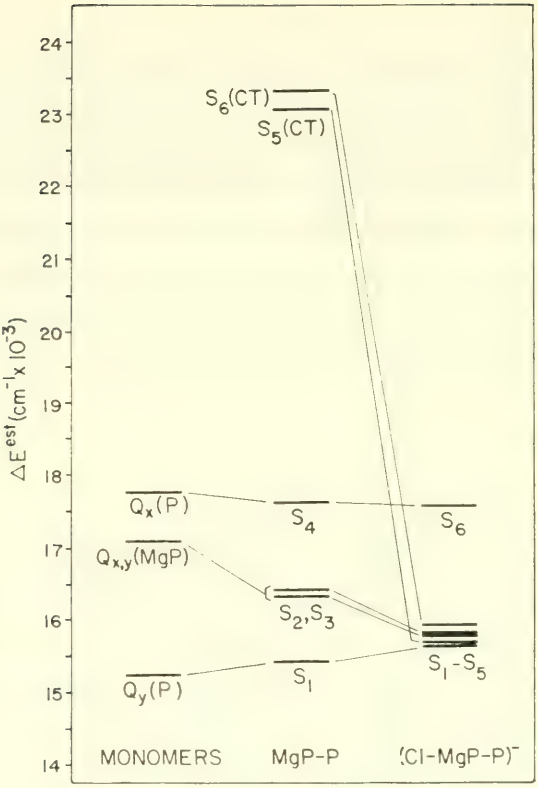


Figure 23. The relationship of gas-phase P-MgP CT states to the excited singlet (π,π^*) states of its monomeric components is depicted in the figure for dimer geometry III. Charge-transfer states S_5(CT) and S_6(CT), lie considerably higher in energy than the low-lying "monomer-like" singlet states $S_1 - S_4$, and due to their presumably rapid rate of radiationless relaxation to the lower excited singlet states, are not suitable candidates as "photo-trap" states which will likely lead to further, irreversible electron transfer. As described in the text, the energies of the highly polar CT states can be dramatically lowered by a variety of microenvironmental effects such as solvent polarity and the presence of anionic species which can bind directly to the dimer complex. This latter effect is also illustrated in the figure, where the addition of the chloride ion is shown to lead to a dramatic lowering of the energies of the two CT states. The relative locations of the CT and (π,π^*) states are similar. In the other geometries investigated, the CT states lie higher above the lowest (π,π^*) state in the gas phase than in geometries II and III, and therefore these dimers appear to be less favorable candidates as photo-traps.

highest occupied molecular orbital of MgP, 1_{MgP}, to the lowest and next lowest unoccupied molecular orbitals of P, 1_P^* and 2_P^*, respectively. Radiative transitions, $S_5(CT) \leftarrow S_0$ and $S_6(CT) \leftarrow S_0$ are predicted to be forbidden. Obviously, in all cases $S_5(CT)$ lies too high in energy for the dimer to act as effective phototrap.

P-MgP In Solution

Netzel et al. (loc. cit.) have studied the picosecond excited state photodynamics of "face-to-face" P-MgP dimers in which the two macrocycles are linked by four- and five-atom chains, corresponding approximately to our geometries II and III. In tetrahydrofuran (THF) solution, they found no evidence of formation of any CT states upon excitation. However, in CH_2Cl_2 they detected formation of a CT state below the lowest (π,π^*) state, which decayed both radiationlessly and through the triplet manifold, as shown in Figure 24. Apparently, a polarizable solvent such as CH_2Cl_2 is capable of considerable stabilization of the CT states, which possess large dipole moments due to their radical-pair character (μ = 24 and 29 Debye for geometries II and III, respectively).

P-MgP•••Cl$^-$ Dimers

Netzel et al. (loc. cit.) also performed experiments on P-MgP in $CH_2Cl_2/((C_2H_5)_4NCl)$, in which the chloride ion is known to coordinate with the magnesium ion of MgP. They detected formation of a CT state which decayed only by internal conversion to S_0, as shown in Figure 25. This implied that

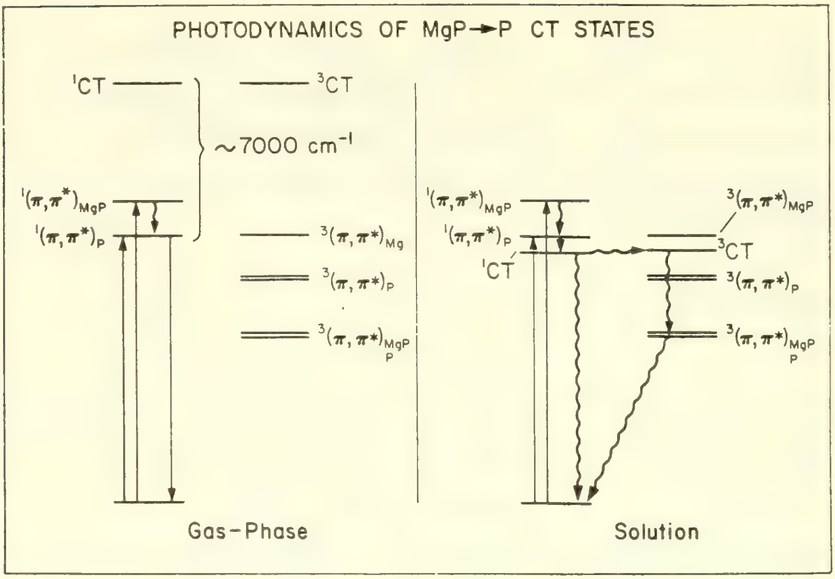


Figure 24. Schematic representation of the photodynamics of MgP → P CT states based on experimental studies of Netzel et al., and on theoretical calculations carried out in the present work. The figure illustrates the influence of solvent polarity on the overall photodynamics. Of particular interest is the enhanced intersystem crossing from ^{1}CT to any of several approximately isoenergetic tripled (π,π^*) states in solution (N.B. intersystem crossing between CT states is forbidden, but can occur through vibronic interactions with nearby triplet (π,π^*) states), which give rise to new pathways for radiationless de-excitation to the ground state thus quenching the fluorecent emission observed in non-polar solvents and presumed to occur in the gas-phase.

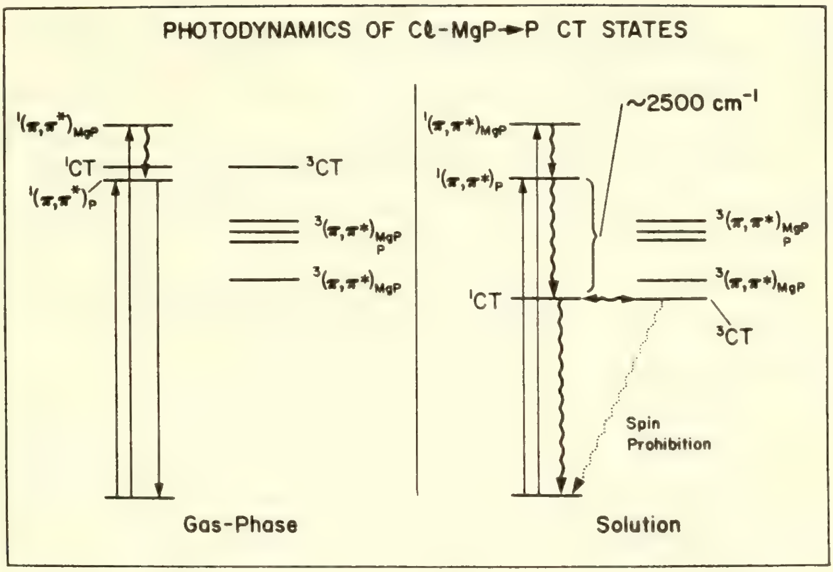


Figure 25. Schematic representation of the photodynamics of $Cl^-\cdot\cdot MgP \rightarrow P$ CT states based on experimental studies of Netzel et al., and on theoretical calculations carried out in the present work. The figure illustrates the additional significant stabilization brought about by the presence of chloride ion. In addition, it should be noted that intersystem crossing between CT states is forbidden, and thus unless vibronic interactions with nearby triplet (π,π^*) states is significant it would be unlikely that the triplet manifold would play a significant role in the overall photodynamics. And as is clear from the figure, this would lead to enhanced depopulation of the lowest singlet CT state in solution via radiationless relaxation to the ground state.

the CT state was lower in energy than the lowest triplet (π,π^*) state of the system.

Figure 23 shows the results of calculations on P-MgP$\cdots$Cl$^-$ (in geometry III), in which a bare chloride ion is coordinated to the magnesium ion at a distance of 3.10 A. Clearly, the CT states are lowered substantially, while the (π,π^*) states are not significantly perturbed. Similar results are observed from geometry II. For other geometries, however, the CT states remain above the lowest singlet (π,π^*) state.

The effect of chloride ion on the molecular orbitals of P-MgP is strictly a field effect which destabilizes each molecular orbital, but does not alter the charge distribution. Consequently, in the following equation which approximately describes the CT transition energy, ΔE, the coulombic repulsion between the 1_{MgP} and 1_P^* molecular orbitals, denoted by J, remains constant upon addition of chloride ion, but the values of the corresponding molecular orbital energies, $\varepsilon(1_P^*)$ and $\varepsilon(1_{MgP})$, are raised.

$$\Delta E = \varepsilon(1_P^*) - \varepsilon(1_{MgP}) - J$$

However, since the chloride ion influences the MgP molecular orbitals more strongly than those on P, the orbital energy difference $\varepsilon(1_P^*) - \varepsilon(1_{MgP})$ diminishes due to the presence of the chloride ion, and the CT transition energy is consequently lowered.

In Netzel's experiment, the effect of the chloride ion is likely to be weaker than in our calculations due to the presence of the counterion. Nevertheless, the combination of solvent stabilization and chloride ion perturbation is apparently sufficient to produce the effect depicted in Figure 25.

Summary of Preliminary Phototrap Design Principles

We have examined isolated P-MgP complexes in a variety of molecular geometries, and have shown that the lowest intramolecular CT state is much too high in energy for the complex to act as a phototrap. It is not expected that relaxation of the CT state to its zero-point vibrational level will alter this picture. Therefore, to trap excitation energy effectively it is necessary that the molecular environment provide sufficient stabilization of the lowest CT state to position it below the lowest (π,π^*) state of the complex.

Given the approximate CT transition energy formula above, it was shown that geometry variation only influences the value of the intermolecular repulsion J, and therefore geometries in which π-π interactions between the macrocycles are larger, will have lower CT states. However, dimers in which the two rings are more widely separated will have CT states with larger dipole moments, and thus may be more effectively stabilized by polar or polarizable solvents.

Within a fixed geometry, ΔE may be lowered by a) lowering $\varepsilon(1_P^*)$ to make P a better electron acceptor, b) raising $\varepsilon(1_{MgP})$ to make MgP a better donor, or c) both a) and b). We have seen that the presence of chloride ion, acting as

acting as an external charge, is effective in lowering ΔE. These observations suggest the general picture of a phototrap given in Figure 26. For a fixed dimer geometry, the position of the lowest CT state could in principle be adjusted through synthetic modifications of the substituents containing the charged groups and/or groups which are strongly electron donating or withdrawing.

On-Going Studies

Studies in progress are directed towards extending the capabilities of our programs to enable treatment of larger molecular systems. This has necessitated the development and characterization of several new procedures designed to provide computational alternatives to the more "brute-force" approach used in the studies described above. Preliminary calculations on the anthraquinone containing trimer AQ-P-MgP trimer depicted in Figure 27, demonstrate the computational feasibility of carrying out *ab initio* quantum mechanical calculations on such large systems.

Future Directions

While the studies described above have provided a number of useful insights into the electronic and geometric structural features of porhyrin heterodimers as a putative model of a biomimetic solar energy conversion system, the results obtained to date are qualitative or semi-quantitative at best. In order to fully characterize these and related systems, such as the AQ-P-MgP system, more detailed theoretical studies must be undertaken. These studies include improvement of the level of the quantum mechanical description

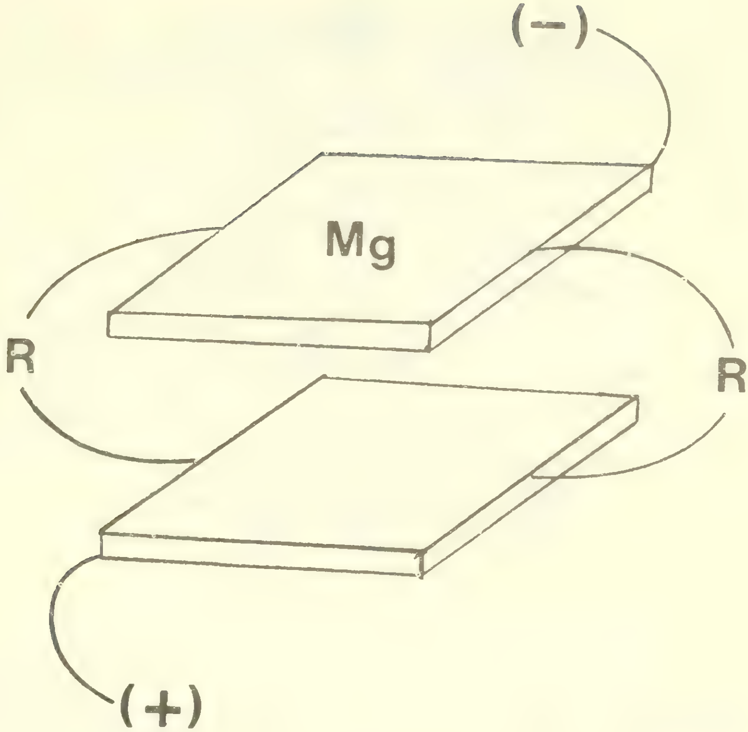


Figure 26. Schematic representation of a model system, based on the asymmetric porphyrin dimer system, designed to produce low-lying CT states that will be able to initiate irreversible, photo-induced electron transfer in an analogous fashion to that which occurs in photosynthetic systems. The two inter-monomer links labeled with R depict the influence of inter-monomer geometry, while the positive and negative charges "linked" to the dimer framework depict an idealized charge configuration that would further enhance the stability of inter-monomer charge-transfer states. These charged entities could be replaced effectively by molecular species which are strongly electron donating or withdrawing.

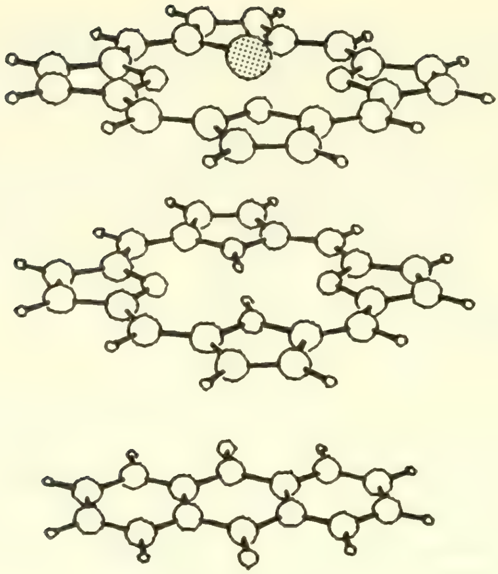


Figure 27. Scale ("ball-and-stick") model of the AQ-P-MgP trimer studied in the present work. The MgP moiety lies at the top of the figure, the AQ lies at the bottom, and the P moiety is sandwiched in between. The total system, which is compromised of ninety-nine atoms and 442 electrons, represents a significant computational endeavor which could not be undertaken without the use of significant computer power such as is available on a Class VI level CRAY 1S or equivalent machine.

through the use of higher level basis sets (e.g. double zeta or better quality), investigation of the vibrational properties of ground and excited electronic states, and inclusion, implicitly or explicitly, of microenvironmental effects. Information provided by these studies is essential if a detailed quantitative picture of such important physico-chemical processes as electronic energy transfer, photo-induced electron transfer, and radiationless transitions, which underlie both biological and biomimetic solar energy conversion, is to be obtained. Calculations of the type described above on such large systems will require even more powerful computers than the Class VI CRAY 1S's used in the calculations described earlier.

Monte Carlo Methods in Quantum Mechanics

An algorithm which will need the capability of Class VII machines is the exploration of the "mirror potential" method of solving non-relativistic many fermion problems in quantum chemistry, nuclear physics and condensed matter physics. To generate an axisymmetric solution by Monte Carlo methods, one writes the wave function as

$$\psi(R) = \psi^{+}(R) - \psi^{-}(R)$$

and introduces a pair of coupled Schrodinger equations

$$H\ \psi^{+}(R) + C(R)\psi^{+}\ (R) = E\psi^{+}(R)$$

where H is the original Hamiltonian and C(R) is an arbitrary symmetric function. This formalism permits the Monte Carlo solution without any fixed

node or other approximation. It requires the simulation of coupled random walks (each in the full coordinate space) since the "potential" for ψ^+ now contains $C(R)\psi^-$ (the "mirror" potential) and vice versa. Thus the walk for ψ^+ is done using an ensemble of walkers for ψ^- to determine the effective potential. We estimate that rather dense sets comprising about 10 million words will be needed to explore this algorithm in an experimental way. The expectation is that such experiments will lead the way to efficient algorithms but it has proved necessary in the past to carry out substantial experiments first. The present work in which approximate mirror potentials are used show also that substantial computing power will be needed: a single experiment (on a nuclear four body problem) requires about one hour of Cray 1 time. It is estimated that hundreds of hours of class VII time will be needed for this research.

Figure 28 shows a recent calculation of the momentum density in liquid ^{3}He presented by P.A. Whitlock and R.M. Panoff at the Los Alamos Conference on "High Energy Excitations in Condensed Matter", February 13-15, 1984. These are based on exact integration of the many-body Schrodinger equation using modern He-He potentials; the agreement with experiment is very good. Approximate theory based on HNC integral-equation treatment of variational wave functions is in poorer agreement.

These calculations will have to be repeated and extended especially for crystal phases to compare with experimental data now emerging. The liquid phase results clearly show an exponential tail at high momenta; this is not present in the crystal phase. We will need to carry out calculations on larger simulation systems for the HCP phase (the true zero temperature ground

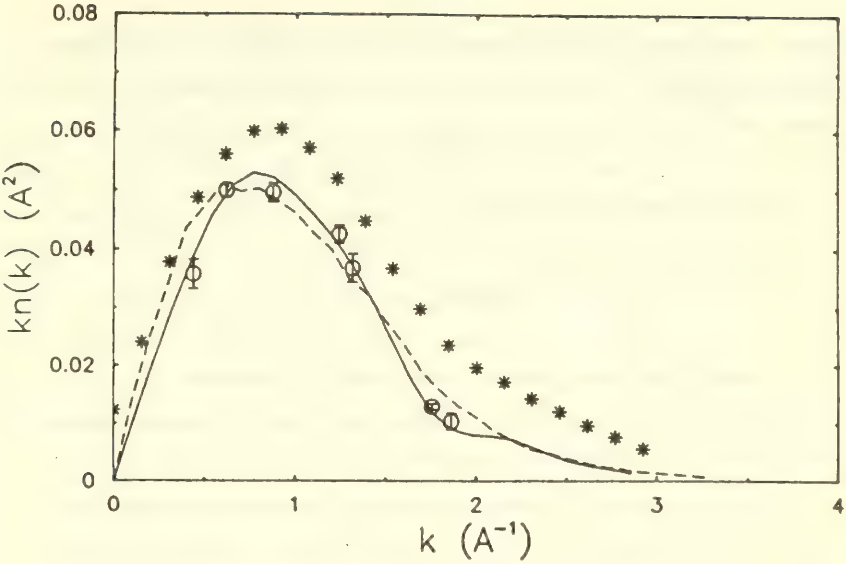


Figure 28. Momentum density for liquid ^{4}He, normalized so that $\int n(k)dk = 1$. The solid line shows kn(k) from our Green's Function Monte Carlo quadrature of the many-body Schrödinger equation. The circles with error bars are a recalculation at discrete momenta in the simulation cube. The dashed curve represents the results from neutron scattering experiments at Chalk River [A.D.B. Woods and V.F. Sears, Phys. Rev. Lett. 39, 415 (1977)]. The stars show the results of hypernetted chain (HNC/0) integral equation theory [M.L. Ristig, Course LXXIX, International School of Physics, "Enrico Fermi", Ed. A. Molinari, North Holland, Amsterdam 1981].

state) using translationally invariant importance functions. These require class VI or class VII computers since they are heavily CPU bound; class VII memory sizes are helpful but not essential for these calculations.

4.3 APPLIED MATHEMATICAL SCIENCES

Nonlinear Dynamics

In the last ten years there has been a revival in nonlinear dynamics after almost a century of dormancy. Much progress has been made, but still, the subject is not in good shape and there is much to be done and understood. The driving force behind the recent developments in nonlinear dynamics is a series of questions posed by theoretical and practical problems in various fields of science and technology. They are found, for instance, in:

fluid dynamics--problem of turbulence

biology--studies of bacterial contamination

ecology--animal dynamics

meteorology--weather forecasting

optics--laser systems

electrodynamics--nonlinear circuits

physiology--dynamics of the heart

economics--market oscillations

chemistry--rate equations

accelerator technology--beam stability

thermonuclear fusion--plasma confinement

celestial mechanics--stability of solar system

oceanography--internal waves, etc.

The common element among these examples is that they are nonlinear systems and can exhibit very complicated motion. Although these systems obey strictly deterministic laws, they seem to behave in a random manner; some properties of their trajectories are as chaotic as the toss of a coin. From the examples mentioned above, chaotic time-dependence is not an exceptional feature of a particular system but a property shared by a broad class of typical systems found in many fields of science and technology. The work in the mathematical aspects of nonlinear dynamics tries to establish basic principles so that scientists and engineers can then apply these principles to understand and analyze the systems they are investigating in their own fields.

This research is nonlinear in nature and hence it relies heavily on numerical experiments. Computers like the CRAY-1 have permitted the study of a wide variety of phenomena in chaotic dynamics. However, this work has been limited to very simple mathematical models and to systems with few dimensions because of the lack of more capable computers with large memories. Two important basic problems which illustrate the need for more powerful computers than the CRAY-1 in the study of nonlinear dynamics are now considered.

<u>Strange Attractors</u>

One important aspect of nonlinear dynamics is the study of the long term behavior of systems of ordinary differential equations. These equations are used to model a physical process, and describe how a given quantity is changing in time as a function of certain other quantities. For instance, these equations might be used to predict next year's mosquito population based on parameters like this year's rainfall, number of birds, acres of swampland,

etc. Knowledge of the solutions of such equations can lead to a more complete understanding of the physical system which they model.

In recent years many nonlinear systems have been discovered whose solutions are chaotic; that is, the solutions oscillate irregularly, never settling down to a regular pattern. Even so, their behavior is not completely random; the solutions approach marvelously complex sets, called "strange attractors", that have remarkable structure. No matter how much one magnifies a given part of a strange attractor, one sees well defined structure as shown in Figures 29 and 30. Examples of chaotic systems of equations have been found in disciplines as diverse as solid state physics, neurophysiology, meteorology, and plasma fusion as we have mentioned earlier. Although the equations have widely different forms, the ways in which the solutions become chaotic and the properties of the chaotic solutions themselves seem to exhibit certain patterns. It now appears that chaos is a fundamental property of nature. Just as the theory of relativity revolutionized physics in the early part of this century, a mathematical understanding of chaos and strange attractors is of enormous scientific importance in a wide variety of fields.

Another universal property of chaotic systems is their sensitive dependence on initial conditions. Two distinct initial states, no matter how close together, will produce trajectories that stay close together only for a short time, then separate exponentially quickly. This exponential separation of nearby points means that a chaotic system "loses" information at a certain rate. In practice, this means that if one measures an initial state experimentally, one can predict the behavior of a chaotic system only for short periods. No matter how accurate the measurement, there is always some

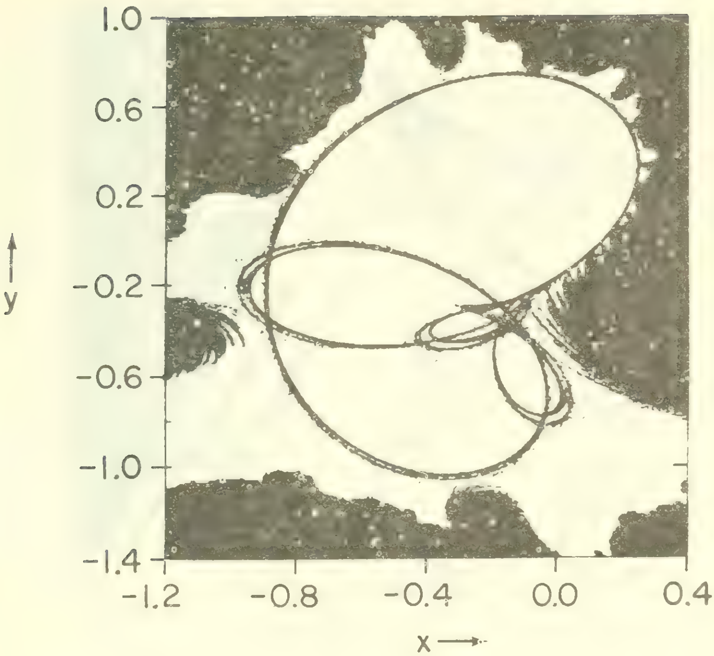


Figure 29. Picture of a two-dimensional strange attractor shown by the convoluted lines in the blank region. Initial states in the blank region are attracted to this strange attractor while initial states in the black region are attracted to some other attractor.

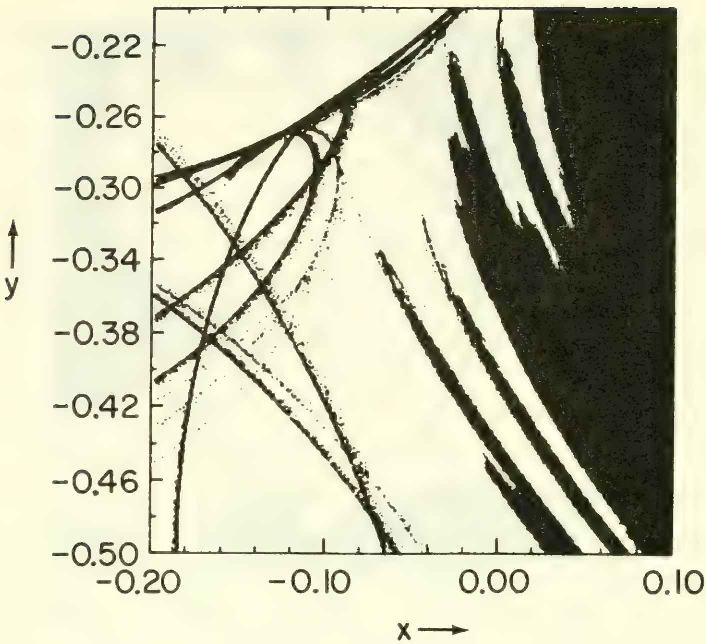


Figure 30. This picture is a blow up of a small region of Fig. 29. We can see that the convoluted strange attractor exhibits a fine scale structure which will always be present, even after repeated magnifications.

experimental error, so that the "true" initial state is different from the measured one. Although the equations that model the system are deterministic, the experimenter will not be able to predict the long-term behavior of the process, because the error between the "true" and the measured initial states will be magnified until the trajectories which start from them appear to evolve independently.

These considerations lead to the following questions. Can one measure the rate of separation of nearby points? Is it possible to characterize the structure of a strange attractor in some systematic way? Mathematicians have studied notions of "dimension" of an attractor to try to answer these questions. It appears that the various definitions of dimension, discussed below, can distinguish a chaotic from a non-chaotic system, assess the system's sensitivity to initial conditions, and also give a lower bound on the number of essential variables needed to model the dynamics. This is not an arcane consideration; a real-world model may have dozens of input variables, and a researcher may need to know not so much as whether the dimension is 3.1 or 3.2 as whether the dimension (and therefore the number of variables which affect his system the most) is 3 or 30. A suitable notion of dimension, then, can yield a basic qualitative understanding of the dynamics of a given system. The goal of this research is to establish easily-applied mathematical principles that can help other scientists make sense of a wide range of complex behavior in many natural systems.

The definitions of dimension have been motivated by other branches of mathematics, including information theory. Because the solution of initial-value problems separate exponentially quickly, the system "loses"

initial information, i.e., information acquired about the system many time steps ago is less and less predictive of the current state. Conversely, the chaotic motion of trajectories acts to create new information, since the longer one observes the system, the more one learns about it, because typical trajectories are not periodic, that is, they never do the same thing twice.

Many tools presently available to study this kind of behavior rest on the abstract mathematical idea of a "covering." Suppose one is given a set of points in space and an inexhaustible collection of tiny boxes; one moves the boxes around in such a way that every point of the set is covered by (contained in) some box. Usually one wants to move the boxes around in such a way that the minimum possible number is used to cover the set. There are many results that can be proved in information theory that are based on the idea of covering sets with boxes. By studying the covering of a set of points, one can gain information about the way points are distributed within the set. This is important in the study of strange attractors since these sets have extremely complicated structure. One can also use the results of covering theorems to say something about how much information is gained by taking periodic "snapshots" or experimental measurements of a system. It is generally believed that these ideas can be applied profitably to the study of chaotic dynamics.

It is possible to calculate quantities known as the Lyapunov numbers for a dynamical process with relative ease. These numbers indicate whether a system is chaotic, and if it is, how quickly trajectories starting from nearby initial conditions will separate. More important, there are several conjectures which relate the Lyapunov numbers to the other, "covering-based"

definitions of dimension. Since Lyapunov numbers can be computed readily on a minicomputer, these conjectures have sparked a great deal of interest. Consequently, one needs to determine how well the conjectures are satisfied for a variety of systems.

Although a powerful abstract idea, it is extremely difficult to generate a covering of a set in as few as two or three dimensions on a computer, yet such low-dimensional cases are fundamental to an understanding of the structure of chaotic systems. To appreciate the problems involved, imagine a three-dimensional cube whose sides are each one unit in length (say one inch). Suppose that inside this cube lies a strange attractor for which one wants to find a covering. To make reasonable measurements of the dimension of the attractor, one needs a covering by very small boxes. If one takes the boxes to be 1/100 unit in length (a typical size for such a measurement), then potentially 1,000,000 boxes need to be considered just to find a covering for the attractor! Unfortunately, such a procedure taxes the memory capacity of the present Cray-1 computers, yet this kind of measurement is precisely what is needed to begin to gain a mathematical understanding of chaotic systems in as few as two or three variables. However, more capable machines, such as a Class VII computer, bring such a calculation into the realm of feasibility.

Other kinds of dimension calculations do not require a great deal of memory. However, it is necessary to follow a group of trajectories for a long period of time; it is not unusual to generate 10,000,000,000 trajectories to make a reasonable estimate of a dimension. Again, even on a pipelined, vector processing machine like the Cray-1, such a calculation, though theoretically

possible, takes too long to do on a systematic basis. However, this kind of computation is well-suited to a multiprocessing environment.

Enhanced supercomputer capability and capacity is essential to this research because it is necessary to compare the results from set-covering algorithms with the Lyapunov exponents for a large variety of systems. Numerical experimentation is an indispensible tool in the discovery of new theorems in this field. If the conjectures which relate the two types of quantities can be proven, then the future study of many dynamical processes will be speeded greatly, because researchers can find out many important properties of their systems from the easily-computed Lyapunov numbers.

Basins of Attraction

Another important aspect of nonlinear dynamics is concerned with predictability in science. In other words, given a nonlinear dynamical system and some initial state one would like to be able to predict the final asymptotic state of the system. This final asymptotic state which is approached by a trajectory of the system is called an attractor and the set of all initial states which approach some attractor is called the basin of attraction of that attractor.

A common property in nonlinear dynamical systems is the coexistence of more than one attractor and, hence, more than one basin of attraction. Therefore, the determination of the various basins of attraction is important because it gives the location of initial states which approach the various final states. If an experimentalist is able to measure the initial state of

an experiment with enough precision, thus guaranteeing its location in a given basin of attraction, then immediately it would be possible to predict the final state of the experiment which, in this case, would correspond to the attractor of that basin of attraction.

To illustrate this idea, consider the set of initial states of a nonlinear system to be given by the area of the rectangle schematically depicted in Figure 31. There are two possible final states denoted by the attractors A and B. The region to the left of the curve Σ is the basin of attraction for the attractor A while the region to the right of Σ is the basin of attraction for B. The curve Σ separating both basins of attraction is called the basin boundary. Points 1 and 2 represent two initial states where some error ε, indicated by the circles, is made in measuring the initial state of the system. Since the circle surrounding 1 is totally contained in the basin of attraction for B the trajectory generated by the initial state 1 is definitely attracted to B. But the initial state 2 is uncertain that it may be attracted to either A or B because the basin boundary Σ passes through the circle surrounding 2. However, in this case we can easily resolve the uncertainty by increasing the precision with which the initial condition 2 is measured such that the smaller circle would be totally contained in one of the basins of attraction. This nice situation occurs whenever the nonlinear system has a smooth basin boundary of integer dimension. An actual example of a nonlinear system exhibiting a basin boundary with integer dimension is shown in Figure 32. The blank region is the basin of attraction for the strange attractor which consists of four thin line segments as we can see in the picture. The black region is the basin of attraction for another strange attractor. Although the two basins of attraction are convoluted, the basin is

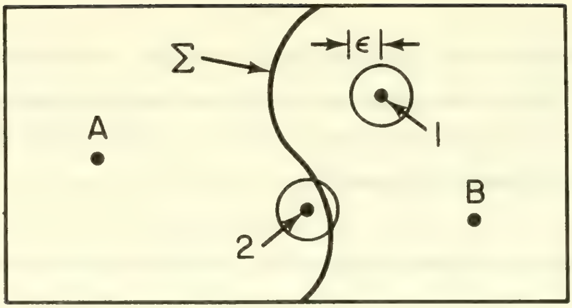


Figure 31. Pictorial representation of the attractors A and B. The region to the left of the curve Σ is the basin of attraction for the attractor A while the region to the right is the basin of attraction for B. The points, 1 and 2, indicate initial states measured with some error ε given by the small circles. The initial state 1 approaches the attractor B while the initial state 2 is uncertain in the sense that it can go to either attractor A or B.

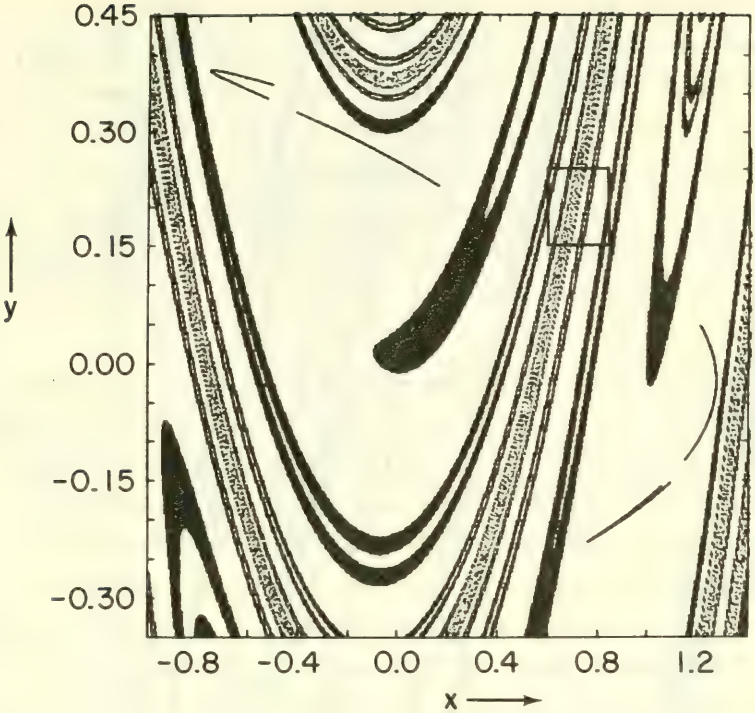


Figure 32. Example of a nonlinear dynamical system exhibiting multiple strange attractors. The blank region is the basin of attraction for the strange attractor which consists of the four thin line segments. The black region is the basin of attraction for another strange attractor. The basin boundary, which is the curve separating both basins is a smooth curve of dimension 1.

a smooth curve of dimension 1. In the same figure a region is boxed in which the basins of attraction look complicated. However, Figure 33 shows part of the blow up of that boxed region by a factor of 25 and, indeed, the basin boundary is one dimensional and further magnifications would show the same structure. For this kind of basin boundary, the accuracy of prediction of the outcome of an experiment is directly proportional to the precision with which the initial states are measured.

Dynamical systems may have much more complicated basins of attraction. Figure 34 shows the basins of attraction of an example of this kind of system. Initial conditions in the black region approach the point attractor indicated by an arrow on the left while initial conditions in the blank region yield solutions which are asymptotic to the other attractor indicated by the arrow on the right of the picture. For this system, the basins of attraction are exceedingly intertwined. A blow up of a tiny region of Figure 34 by a factor of 10^9 would show the same finely intertwined structure of the basins of attraction. The basin boundary has a noninteger dimension equal to 1.8, i.e., it is a fractal curve. For this nonlinear system, even if the precision with which the initial state is measured is increased, the final state would still be uncertain because we would not able to tell in which basin of attraction the initial state is in. In Figure 35, the fraction of initial states $\bar{f}$ which is uncertain versus the error ε in measuring the initial state is plotted. From the graph it is seen that when $\varepsilon = 0.1$ then 60% of the initial states are uncertain. If ε is reduced by a factor of 1,000 to $\varepsilon = 0.0001$, still 20% of the initial states would be uncertain. It means that if an experimentalist is faced with this kind of system, a tremendous improvement in the precision with

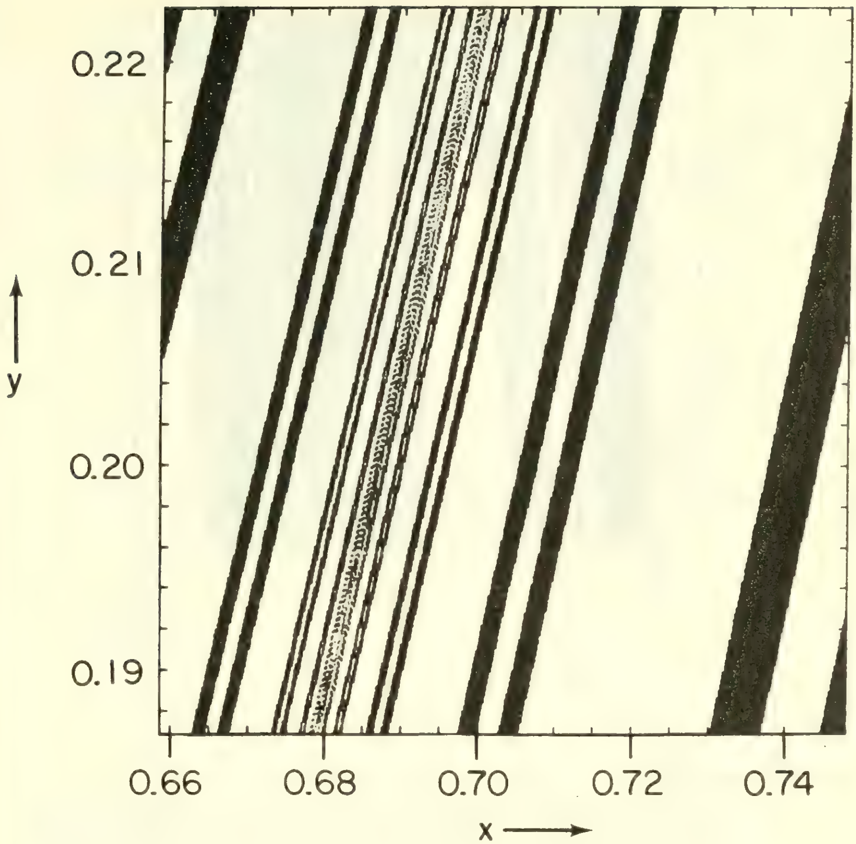


Figure 33. This picture is part of the blow up of the boxed region shown in Fig. 32 by a factor of 25. We can see that the dimension of the basin boundary is indeed 1 and further magnifications would show the same structure.

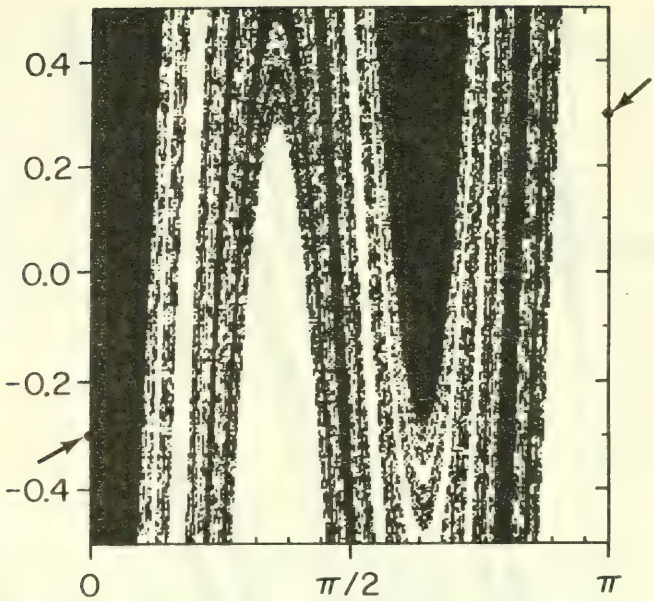


Figure 34. Example of a nonlinear dynamical system exhibiting exceedingly intertwined basins of attraction. The initial conditions in the black region approach the point attractor indicated by an arrow on the left, while initial conditions in the blank region are asymptotic to the other attractor indicated by an arrow on the right. A blow up by any amount of a tiny region of this picture would show the same finely intertwined structure.

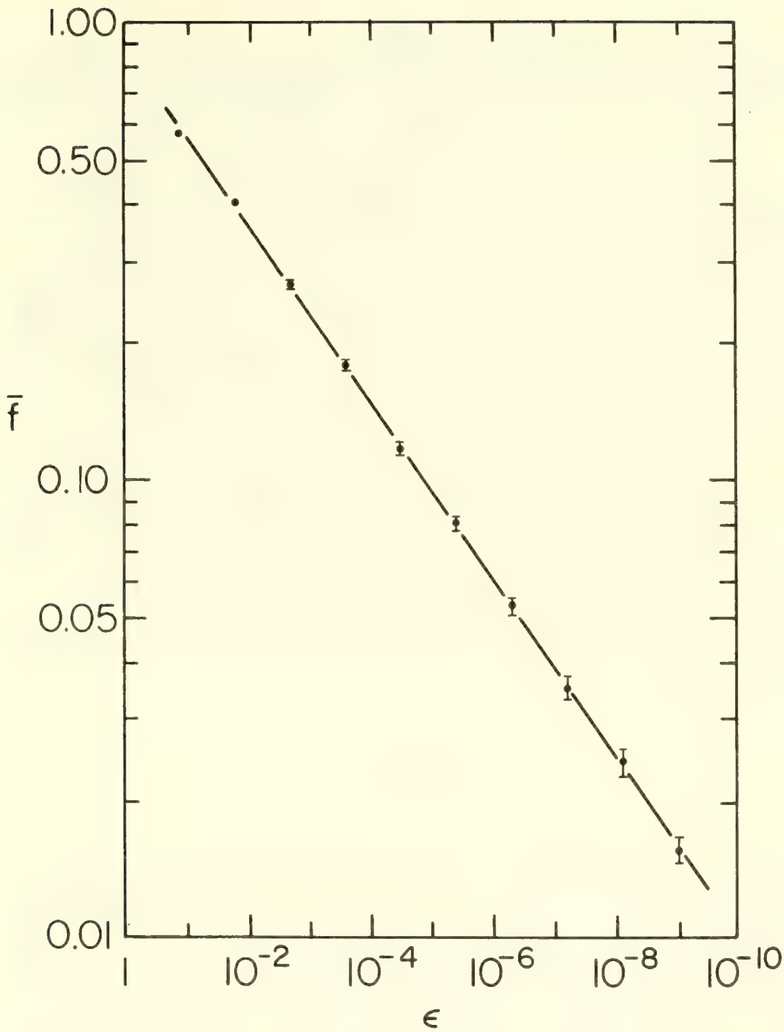


Figure 35. Plot of the fraction of initial states $\overline{f}$ which is uncertain versus the error ε in measuring the initial state. We can see that when $\varepsilon = 0.1$ then 60% of the initial states are uncertain. If we reduce ε by a factor of 1,000, still 20% of the initial states would be uncertain. It means that an experimentalist has to work very hard to improve the precision with which the initial state is measured to have only a modest gain in the capacity to predict the outcome of the experiment.

which the initial state is measured would give only a modest gain in the capacity to predict the final outcome of the experiment.

It is important then to provide the experimentalist with sufficient information about the nonlinear dynamical system that most correctly models the experiment. However, the investigation of the structure of basins of attraction can only be done numerically. Figures 29, 30, 32, 33, and 34 show pictures of basins of attraction of two dimensional systems. To generate each one of these pictures, it is necessary to follow numerically 400,000 trajectories of the nonlinear system. Even with a fully vectorized code, it takes 22 minutes of Cray-1 time. But there is no reason for a given experiment to be adequately modeled by a one- or two-dimensional system. For a three-dimensional system, each run would take more than 50 hours of a Class VI computer. It is easy then to conclude that as the dimension of the nonlinear dynamical system increases we need computers with the capability of a Class VII machine and beyond.

Computer Science

Much of the Applied Mathematical Sciences (AMS) program is concerned with the invention and development of mathematics and computer science techniques that can be used in the solution of problems in science and engineering. To have the greatest potential for benefiting DOE's program, the results of this research must be applicable to realistic problems and must be effective on the leading edge computers in use by DOE research programs. Therefore, AMS investigators must have access to the most up to date and capable computer systems. While the development of effective techniques for the new machines

may require only modest amounts of time on these computers, experimentation with their capabilities and architecture are essential.

The Computational Mathematics part of the AMS program concerns itself with the development of algorithms and software for the frequently occurring computations that are building blocks of many numerical simulation codes. Solution of systems of linear and nonlinear equations, numerical integration, and function approximation are examples of such computations. Today's Class VI machines feature vector architecture. CRAY-1, the most widely used, has been in use for over seven years, yet better algorithms and programming techniques are still being developed for them. In some fields, notably numerical optimization, there are still virtually no algorithms that are tailored to vector architecture. Enhanced Class VI and Class VII machines will feature a few (2 to 16) parallel processors, each of which will be vector oriented. Even this architecture is novel and there is little body of experience to guide its use. Algorithm developers must run experiments on real machines with these features to create effective and efficient algorithms for them.

As a case in point, consider algorithms and software for solving systems of linear equations. These computations are at the core of most numerical modeling work with an estimated 70% of all computer cycles used in scientific computation devoted to linear system calculations. The reason for such a high percentage is that the solution techniques for solving the partial differential equations used to model natural phenomena eventually yield systems of linear equations to be solved.

The LINPACK collection of software is a highly portable and widely used set of Fortran routines that can be used to solve linear equations and perform related computations. At the lowest level of LINPACK are routines known as Basic Linear Algebra Subroutines (BLAS) that perform basic computations like inner products. Given that LINPACK deals almost exclusively with vectors and matrices, it was expected that LINPACK routines would perform well on Class VI computers since they are vector machines. However, typical performance was only about a third of what was expected even with subroutines that were hand-coded to exploit the vector hardware of the CRAY-1.

Analyses of the disappointing performance yield a recasting of the lower level routines so that they concentrate on matrix-vector data structures instead of vector-vector. With no changes to the rest of LINPACK, one can now achieve a much higher level of performance on the CRAY-1. In addition, the new formulation makes it possible to exploit parallelism easily. Preliminary work with a two processor CRAY XMP has yielded speed increases of 1.9 over the one processor performance.

Other parts of the DOE/AMS program are concerned with programming languages, software development and debugging environments, etc. These projects must also gain access to state-of-the-art supercomputers, particularly those with multiple processors that can be controlled by one user program. While language issues for vector machines are reasonably well understood, there is essentially no experience with the use of parallel computers. Even computers with only two or four processors, such as the enhanced Class VI and Class VII machines, can begin to shed light on programming and debugging issues.

As we move into multiprocessor environments, even the mathematical models and discretization techniques may need to change. Thus the applied analysis segment of AMS research will also require enough access to parallel machines to validate the new models and techniques.

In summary, the new supercomputers have architectural features that must be taken into account for their effective use. The DOE/AMS efforts to create general-purpose mathematics and computer science techniques must address these features to be relevant and vigorous, and access to state-of-the-art supercomputers is essential.

Automated Reasoning

Automated reasoning is the study of concepts and techniques that enable a computer program to carry out tasks that have traditionally been thought of as requiring logical reasoning rather than numerical computation. There are three classes of specific research goals.

The first is to deepen our understanding of the components of the reasoning process. These have been broadly classified as representation of knowledge, rules for deducing new knowledge from existing knowledge, and strategies for focusing exploration and expansion of the knowledge base on the problem at hand. The second research goal is to develop automated reasoning software to take advantage of the latest in computer technology in order to verify that our understanding of the reasoning process can be implemented in

computer programs that solve a continually expanding class of problems with acceptable speed. The third is to explore, and to help other scientists explore, applications of this technology in areas where deep reasoning is required. Such areas include the design of logic circuits, the verification of properties of computer programs, and the diagnosis of problems in complex physical systems such as nuclear power plants. In fact, automated reasoning is likely to become heavily used in the formulation, analysis, and validation of the design of future advanced computers.

Limitations of Current Technology

The third area of research, applications to problems in other areas, suffers from the speed of today's computers. Although automated reasoning technology has been successfully applied to difficult problems, including open questions in mathematics and logic, many seemingly simple problems require vast computer resources. Sometimes strategies can be found that economize on such resources in the attack on a particular problem, but many problems require more resources than are available on today's machines. This is especially true of problems that require very prompt solutions. An example of the latter is real-time assistance with the operation of a nuclear plant. To analyze the behavior of such systems, with but a few dozen components, requires seconds of CPU time on the fastest computers. It has been estimated that practical program verification systems must await two orders of magnitude increase in performance over today's levels. Computational resources needed for these applications are large memories and very great processing power (large number of basic operations performed per second).

Need for Higher Performance Computers

As in the numeric area, it is reasonable to expect that the increase in processing power required for advanced applications will be achieved primarily through multiprocessing on systems with parallel architectures. Automated reasoning algorithms are well suited to such an approach. Many tasks can profitably be done in parallel by a family of cooperating processes, much as a corporation can accomplish tasks too large and complex for a single person. Class VII computers and beyond are expected to provide multiple processing units suitable for the execution of such algorithms.

The availability of much faster computers would expand the scale of problems to which automated reasoning techniques can be applied. This would benefit every application; for example, larger circuits, larger programs, and larger plant subsystems could be studied. A speed increase of three orders of magnitude, for example, may make feasible the real-time control of moderately complex physical systems say with several thousand components monitored. In many prospective applications the scale change is necessary in order for automated reasoning technology to be applied in an economically useful way. For many automated reasoning tasks, little or no floating-point arithmetic is performed. Today's Class VI computers are largely optimized to floating point computations. Therefore, to get two orders of magnitude speed up over even a VAX 11/780, we may have to wait for Class VII machines.

Pure research in automated reasoning would also benefit. Ideas come from trying to solve hard problems. Expanding the class of problems that automated reasoning can successfully attack changes also the nature of the problems that

are beyond its reach. These are the problems that spur theoretical progress. There is no prospect of running out of reasoning tasks; there are problems that will tax theoretical techniques and the fastest computers available for generations to come.

4.4 HEALTH AND ENVIRONMENTAL RESEARCH

Overview

Supercomputers in Health and Environmental Research are important for two clearly defined problem areas at opposite extremes of a linear scale: first, at the molecular level which analyzes macromolecular structure and its perturbation by an invasive chemical or radiation environment to obtain biochemical and finally clinical understanding; the second, of regional geographical scale, analyzes the three-dimensional motion of atmospheric or subsurface liquid/vapor systems to determine and attempt to control the distribution of unwanted byproducts of energy generation.

(i) Macromolecular Structure

The structure and function of proteins and nucleic acids provide the basic- science underpinning for much of present-day biomedical science. Classical studies of three-dimensional structure by X-ray and neutron diffraction remain of importance, and this work requires class VI and higher computers for cases of extreme complexity. Beyond this problem area, the new discipline of macromolecular dynamics, the analysis from first principles of

the coupled motions of thousands of atoms, in globular proteins or long-chain nucleic acids, makes computational demands equal to and beyond any present projection of computer capability. Included in this problem classification is the prediction of equilibrium structure from first principles, in particular protein folding and the prediction of secondary structure for RNA. Rewards from successful solution of these difficult problems will include the ability to design proteins to achieve specific biological function, since the required prerequisite capability for tailoring a DNA sequence (which codes for protein) already exists.

Examination of Individual Problems

(1) Protein dynamics.

The present day view of protein structure is largely static, owing to the importance of X-ray and neutron diffraction in structure determination and the fact that these methods give the average positions of the constituent atoms. However, known sections of the primary sequence of a protein may give no well defined diffraction pattern and are thereby recognized as disordered or in large-amplitude motion.

Functional recognition of substrates by enzyme active sites are known to require dynamic conformational adjustments by the protein. Penetration of oxygen to the heme binding site of myoglobin and hemoglobin absolutely requires fluctuation in globin structure which transiently opens a channel. Electron transfer processes important in photosynthesis are believed to depend

on vibrational coupling which dynamically alters the distance between donor and acceptor.

The relative motion of distinct structural domains of a protein is essential in the function of the muscle protein, myosin and in the action of antibody molecules, and in the assembly of viruses. Crucial activities of the nervous system and of the energy conversion machinery in the cell depend on dynamical opening and closing of protein ion channels through membranes.

Many additional examples could be cited, but we conclude the present list by memtioning that each protein, during synthesis or just after, folds into a perfectly defined 3-dimensional structure. The theoretical prediction of that structure from the primary sequences is, so far, beyond us. If this gap can be closed, then proteins of particular required structure can be synthesized, since the primary sequence is already under control via ab initio chemical DNA synthesis or (presently easier) via site-specific mutagenesis (See Figure 36).

Present dynamical codes require tens of hours of Cray-1 equivalent computing time to simulate a few tens of picoseconds (10^{-12} sec) of motion of the atoms in a moderate-sized protein. During this time period the basic oscillatory and local rotational motions of individual atoms and rigid molecular groups become well established and significant averages and ranges for many dynamical variable can be obtained. However, the important, slower cooperative and transitional motions -- adjustment of an antibody hypervariable region so as to bind its antigen or allosteric conformational transition of an enzyme on binding a control cofactor -- are not usually even approached in this analysis. In part, advances in the theoretical description

and in the detailed algorithms are required, but it is clear that severe lack of computing capability is inhibiting the required innovative approaches to solutions.

(2) RNA secondary structure.

The prediction of secondary structure for RNAs of biological importance is both CP-intensive and demanding of large core memory. Internal bonding occurs not just locally but between distant parts of the long-chain molecule and many alternative structures must be compared quantitatively. Present algorithms require typically several hours of Class VI CP time for the analysis of small- to intermediate-sized molecules (Figure 37). Class VII machines of memory size in the range currently contemplated (see Introduction) should make possible the analysis of any RNA of biological importance at the cost of a few hours of CP time per molecule.

(ii) Atmospheric and subsurface fluid flow and the undesired distribution of health-threatening products of energy generation (Figure 38).

The meteorogical and hydrodynamical systems which contribute the core problems in this area are already well known to require the best computational facilities available. Remaining aspects, such as the partitioning of pollutant chemical species between vapor and liquid phases in the generation of acid rain, intensify the computational demands.

(iii) Other requirements. Unusual special problems such as tomography of complex molecular structures by electron microscopy, or study of drastic local distortion of nucleic acids by metal ions as an explanation for metal-ion toxicity, are examples of need for the <u>capability</u> of Class VI and VII computers, though capacity demands are expected to be modest.

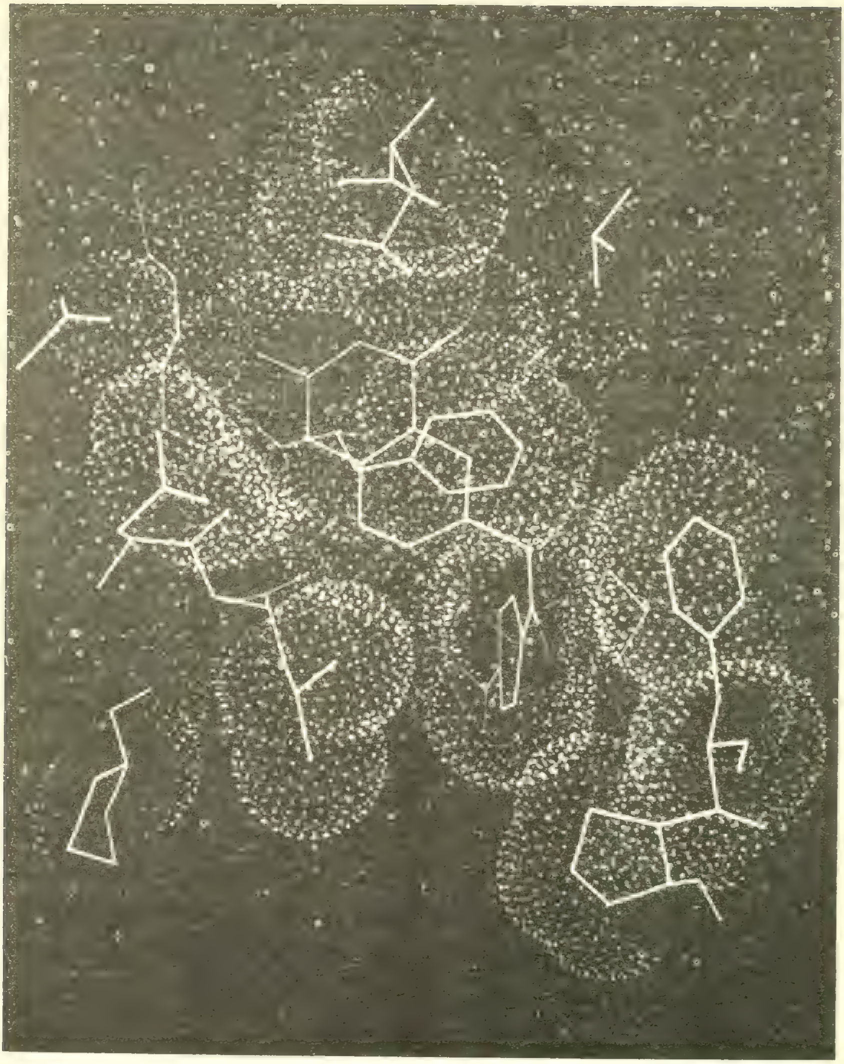


Figure 36. A three-dimensional representation of the anti-cancer drug, Methotrexate, binding to its target site on the surface of the protein Dihydrofolate Reductase. The shapes of the enzyme surface and the binding cleft were obtained from x-ray crystallography. In rational drug design, the drug molecule's shape is tailored to cleft shape. If protein folding could be predicted from the primary sequence, cleft (active site) shape could be altered to affect function, for this and other enzymes.

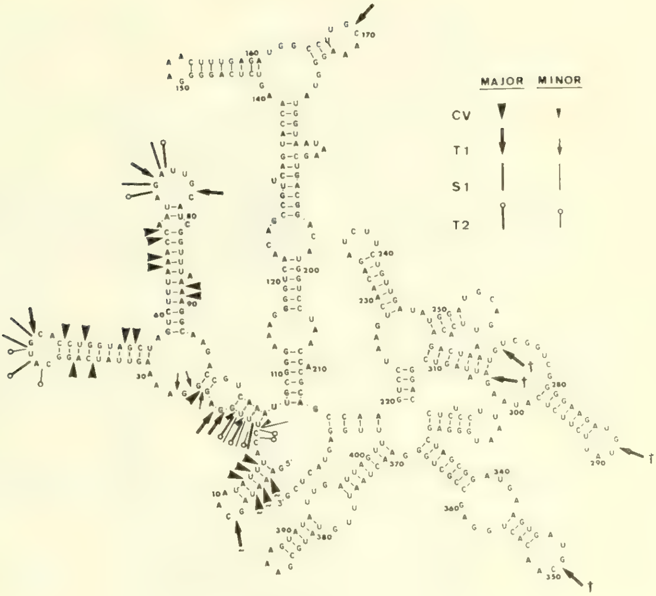


Figure 37. An example of predicted secondary structure for a 414-nucleotide RNA of known sequence from the protozoan Tetrahymena. Note that pairing occurs between stretches of nucleotides which are widely separated in the primary sequence. The arrows show cleavages observed experimentally for nucleases of single- or double-stranded specificity used as tests of the predicted structure.

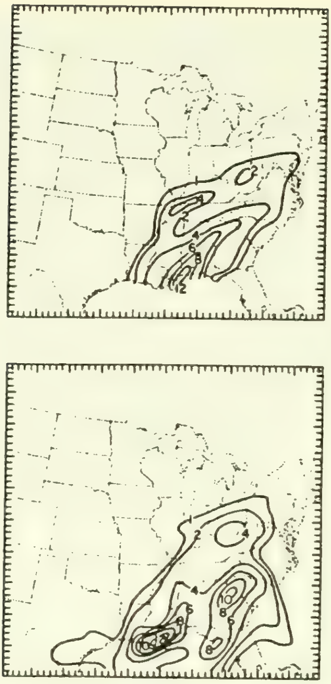


Figure 38. Prediction and control of dispersal of unwanted byproducts of energy generation. Observed (upper panel) and predicted (lower panel) precipitation from a storm system in the southeastern United States in January, 1978. For prediction of acid rain, calculations of SO_2 and SO_4 concentrations in gaseous and aqueous phases of clouds are combined with purely meteorological predictions such as this one.

Selected Bibliography on the Computer Needs for Energy Research Programs

1. "Report of the Panel on Large Scale Computing in Science and Engineering," Peter D. Lax, Chairman, December 26, 1982.

2. "Prospectus for Computational Physics," Report by the Subcommittee on Computational Facilties for Theoretical Research to the Advisory Committee for Physics, Division of Physics, National Science Foundation, March 15, 1981, William H. Press, Chairman.

3. "Future Trends in Condensed Matter Theory and the Role and Support of Computation," report by the Condensed Matter Theory Panel, published under the auspices of the council on Material's Science for the Division of Materials Sciences, U.S. Department of Energy, J. D. Joannopoulos and A. N. Berker, Co-Chairman, September 22, 1981.

4. "An Assessment of Trends in the Solid State Sciences," National Research Council, July 1980.

5. "Future of Nuclear Science," by Ad Hoc Panel on the Future of Nuclear Science, 1977.

6. "Long Range Plan for Nuclear Science," excerpts from a report by the Nuclear Science Advisory Committee, July 30 - August 4, 1979.

7. "Report by the Subcommittee on Computational Capabilities for Nuclear Theory to the DOE/NSF Nuclear Science Advisory Committee," December 1981.

8. "A Report on Computer Needs for Nuclear Theory," William M. MacDonald et al., Seminar Workshop at the University of Maryland, SWUM, April 27, 1979.

9. "Report on the Results of a Questionnaire on Computer Facilities Needs for Nuclear Physics Research," Richard D. Koshel and William M. MacDonald, SWUM, May 14, 1979.

10. "Report of the Sub-Panel on Theoretical Computing of the High Energy Physics Advisory Panel" DOE/ER 0205 September 1984.

11. "Subsurface Transport Program, Geochemical, Hydrological and Biological Cycling of Energy Residuals," Frank J. Wobber, DOE/ER-0156/2 September 1984.

DOE/ER-0240
UC-32

Summaries of the FY 1984 and FY 1985 Applied Mathematical Sciences Research Program

July 1985

U.S. Department of Energy
Office of Energy Research
Scientific Computing Staff
Applied Mathematical Sciences
Washington, D.C. 20545

TABLE OF CONTENTS

TABLE OF CONTENTS continued

PREFACE

The purpose of this report is to provide a convenient compilation of summaries of the individual research projects that constitute the Applied Mathematical Sciences research subprogram managed by the Director, Scientific Computing Staff, Office of the Director of Energy Research, Department of Energy. This new office, organized in 1984, consists of the applied mathematical sciences research activity formerly under the Division of Engineering, Mathematical and Geosciences, Office of Basic Energy Sciences, Office of Energy Research, and the management of the expanded National Magnetic Fusion Energy Computer Network formerly under the Office of Magnetic Fusion Energy. The organization chart on the next page illustrates the structure of the basic research arm of DOE - the Office of Energy Research. The personnel responsible for this subprogram are:

Dr. James F. Decker, ER-2
Deputy Director, Office of Energy Research
Acting Director, Scientific Computing Staff

Dr. Donald M. Austin, ER-7
Applied Mathematical Sciences

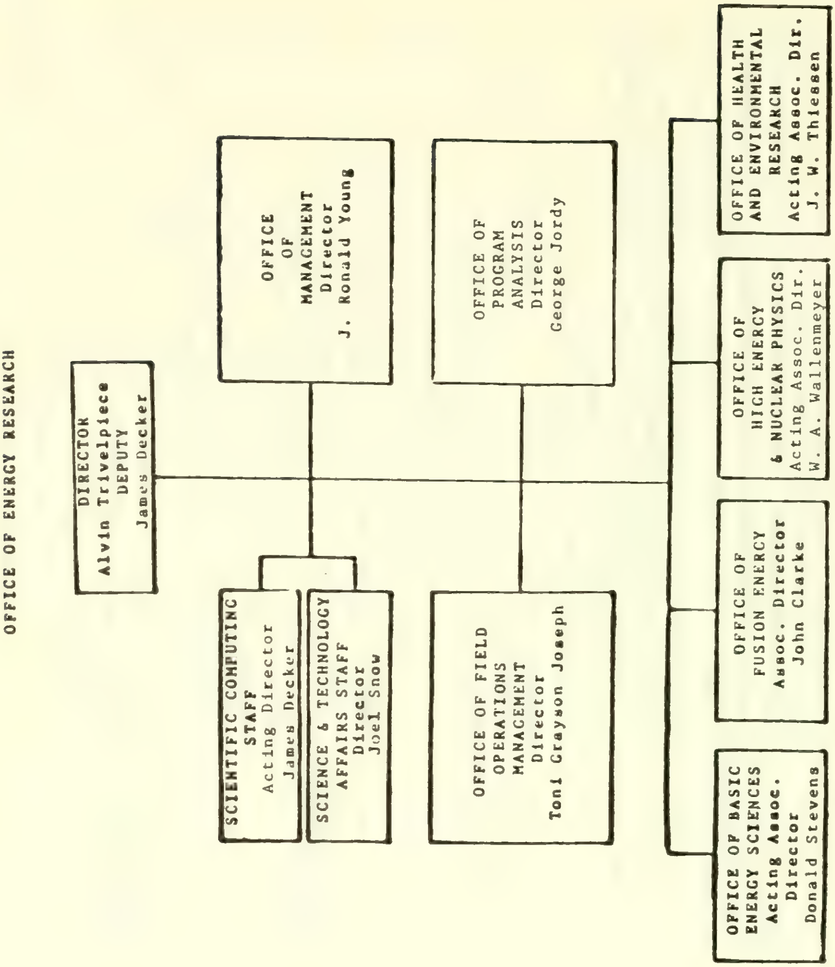
Mr. John Cavallini, ER-7
Energy Sciences Advanced Computation

On temporary assignment to DOE for parts of the period covered in this report are:

Prof. James M. Greenberg, ER-7
Applied Mathematical Sciences

Dr. C. Edward Oliver, ER-7
Applied Mathematical Sciences

Scientific Computing Staff
Office of the Director of Energy Research
U.S. Department of Energy
Washington, DC 20545


OFFICE OF ENERGY RESEARCH
DIRECTOR
Alvin Trivelpiece
DEPUTY
James Decker
OFFICE OF MANAGEMENT
Director
J. Ronald Young
OFFICE OF PROGRAM ANALYSIS
Director
George Jordy
SCIENTIFIC COMPUTING STAFF
Acting Director
James Decker
SCIENCE & TECHNOLOGY AFFAIRS STAFF
Director
Joel Snow
OFFICE OF FIELD OPERATIONS MANAGEMENT
Director
Toni Grayson Joseph
OFFICE OF HEALTH AND ENVIRONMENTAL RESEARCH
Acting Assoc. Dir.
J. W. Thiessen
OFFICE OF HIGH ENERGY & NUCLEAR PHYSICS
Acting Assoc. Dir.
W. A. Wallenmeyer
OFFICE OF FUSION ENERGY
Assoc. Director
John Clarke
OFFICE OF BASIC ENERGY SCIENCES
Acting Assoc. Director
Donald Stevens

Department of Energy
Office of Energy Research
Office of the Director
Scientific Computing Staff
Applied Mathematical Sciences Research Subprogram

Overview

The Applied Mathematical Sciences Research subprogram consists of two parts: Energy Science Advanced Computation and Applied Mathematical Sciences Supercomputing Research. The Energy Sciences Advanced Computation activity provides scientific supercomputer access required by researchers in the Office of Energy Research programs in High Energy and Nuclear Physics, Basic Energy Sciences, and Biological and Environmental Research. This scientific computing will be provided through network access to supercomputers for contractors in universities, industry, and DOE national laboratories. The Supercomputing Research activity supports research in mathematics and computer science that is essential to the use by many DOE programs of the newly emerging multiprocessor supercomputer systems and the long range development of future supercomputer systems. In addition to meeting DOE needs, this subprogram will make a substantial contribution to solving the national problems of providing access to researchers, particularly in universities, and maintaining U.S. leadership in supercomputer technology.

Supercomputers are the fastest general purpose scientific computers available to solve large scale scientific problems. They are today capable of performing 100's of millions of arithmetic operations per second and have several million words of fast memory. The effort in large scale scientific supercomputing is distinguished from the Japanese Fifth Generation project and DARPA's Strategic Computing and Survivability program, which are focused on artificial intelligence techniques.

In recent years, the modern supercomputer has changed the nature of scientific research and technology development. The advancement of science used to depend upon experiments for data and theory for understanding. Today there is a third equally important ingredient to scientific research: computational science. Computational science serves a role that is a hybrid between theory and experiment. In some cases, computations provide insights into experimental data, and in others, computations are used to simulate the ideal experiment to test an analytical model. The emergence of computational science is an important element in

scientific research and technology development is the result of the development of our ability to do computational modeling of complex physical problems and the enormous power of the modern supercomputer. This combination allows scientists and engineers to model complex problems in a much more realistic way and to obtain much more accurate answers than was possible just five years ago.

Scientific supercomputing now makes substantial contributions to many fields of research such as fluid dynamics, plasma dynamics, astrophysics, materials science, chemistry, atomic physics, etc. Similarly, large scale calculations are making significant contributions to product development in areas such as nuclear weapons, electronics, automobiles, aircraft, and chemicals. In product development, computer modeling is used to reduce the number of design/test iterations which are so time consuming and expensive.

The Department of Energy has made a significant commitment to large scale scientific computing in several of its research and development programs. As an agency, DOE is the largest user of supercomputers in the world. In 1985 DOE laboratories will house 22 supercomputers. DOE's commitment to scientific computing is not new. Historically, the Department and its predecessor agencies (the Atomic Energy Commission and the Energy Research and Development Administration) caused many supercomputers to be developed through its nuclear weapons design work at the Lawrence Livermore National Laboratory and the Los Alamos National Laboratory. The supercomputers that were developed in response to DOE needs were delivered to the DOE laboratories nearly devoid of software, requiring our laboratories to develop the necessary software for a complete supercomputer system. Through the software development efforts and the continual interaction between DOE laboratories and the supercomputer vendors, DOE has made significant contributions to supercomputer technologies.

DOE's commitment to scientific supercomputing has paid large dividends. Typical use of supercomputers in the Department's mission oriented development programs is to supplement or replace expensive, time consuming experimental testing with computer simulations. The advantages of such simulations to replace some of the testing in the nuclear weapons development programs are obvious. In a research and development program such as magnetic fusion where experiments cost $10's to $100's millions and take 3-7 years to build, it is impossible to build a new device and test experimentally every good idea that comes along. Computer simulation has become essential.

Great progress has been made in our ability to do computer modeling of complex scientific problems; however, there is still a long way to go. Many of the Department's research and development programs have large, complex research and engineering problems to solve that can be only crudely approximated with today's supercomputers and mathematical techniques. There is a need to include more "physics" in the models, to use finer zoning for more accurate answers and to use more dimensions. In order to solve these complex problems to the accuracy desired in the future, computers with substantially increased computing power over those currently available will be required. Several DOE programs have estimated that they will require computers 200 times as powerful as a Cray-1 or a Cyber 205 by the end of the decade. In fact, some of the DOE programs could use that kind of computing power today, if it were available.

Interagency Coordination of Supercomputer Access and Research Programs

In 1983, the White House Office of Science and Technology Policy (OSTP) established three subcommittees under the Federal Coordinating Council on Science, Engineering, and Technology (FCCSET) to examine the U. S. role in maintaining leadership in the development and use of supercomputer technology. Membership on these subcommittees includes representatives from DOC, DOD (including DARPA, CIA, and NSA), DOE, NASA, NSF, and OSTP. The subcommittees were charged with developing recommendations on the role of the government in maintaining the U. S. lead in supercomputers, on providing access to supercomputers for our researchers, and on coordinating the roles of each agency in the area of research on artificial intelligence and high performance computing. Each of these subcommittees produced reports for OSTP and continues to provide coordination among agencies in these areas.

Two of the subcommittees, one addressing government supercomputer procurement policies and the other addressing access to supercomputers for researchers, were combined into a single subcommittee, chaired by Dr. James Decker, DOE. This subcommittee maintains contact with industry representatives, sponsors panels to investigate questions concerning computer network access to supercomputers and interagency sharing of resources.

The third FCCSET subcommittee is chaired by Dr. Robert Kahn, DARPA, and its charter is to stimulate the exchange of information within the government on high performance symbolic computing and artificial intelligence. During the course of its initial deliberations it became clear that interest in information exchange extended beyond symbolic processing and AI. It was decided to broaden the charter of this subcommittee to include federal research efforts in very high performance scientific and numerical computing and advanced computer research. A report, entitled "Report of the Federal Coordinating Council on Science, Engineering, and Technology Panel on Advanced Computer Research in the Federal Government," June 1985, has been submitted to OSTP.

Another interagency group that has been coordinating research for many years is the Interagency Committee on Extramural Mathematics Programs (ICEMAP). Active members of this committee include representatives from DOE, DOD (AFOSR, ARO, and ONR), and NSF. Occasional participation in the ICEMAP activities include representatives from NIH, NASA, CIA, NSA, and DOC (NBS). In FY 1985, the ICEMAP agencies cooperated in establishing the Board on Mathematical Sciences in the Commission on Physical and Mathematical Sciences of the National Research Council in the National Academy of Sciences.

Energy Sciences Advanced Computation

The Energy Sciences Advanced Computation activity provides OER research access to supercomputers over nationwide networks, including the NMFECC satellite network, the ARPANET, and commercial value-added networks. Although DOE has been a leader in the use of supercomputers in its defense and magnetic fusion programs, there are several research programs that have not had easy access to modern supercomputers. In FY 1985 the use of the National Magnetic Fusion Energy Network was expanded to researchers

supported by the Department's High Energy and Nuclear Physics, Basic Energy Sciences, and Biological and Environmental Research programs. A Cray-XMP/22 was added to the computer capacity at the MFE center at Livermore and expansion of the communications network was begun. In addition, a new supercomputer research institute at Florida State University was initiated at the direction of Congress. A Cyber 205 supercomputer at FSU was installed and linked to the MFE network in early 1985 for use by OER researchers. This activity is described in DOE Report Number DOE/ER-0218, "The Role of Supercomputers in Energy Research Programs," February 1985.

The Research Activity

The primary objective of the research activity is to advance the understanding of the fundamental concepts of mathematics, statistics, and computer science underlying the complex mathematical models of the key physical processes encountered in the Department's research and development programs. In addition, this activity supports investigations of parallel computer architectures that may lead to new approaches to supercomputers. The Supercomputer Research activity provides the primary source of research funding for applied mathematics and computer science research to meet the DOE needs in this area.

One of the primary goals of this activity is to develop our ability to use multiprocessor computer systems. The multiprocessor direction for supercomputers is dictated by the fact that only very limited gains can be made from increased component speeds and levels of integration in the future. This has caused computer architects to look to using more than a single processor working on a problem. In the near term, the supercomputer vendors are producing systems with 2-16 processors. In the longer term, systems with 100's to 1000's of processors are expected.

The change from von Neumann (sequential) to parallel processor machines represents an enormous change in scientific computing, greater than the challenge faced by the introduction of vector processors. The challenge with parallel processors is to have a large number of processors efficiently working simultaneously on the same problem. New software, languages, algorithms, and even

entirely new mathematical approaches to solving important classes of scientific problems be will required. One of the objectives of this activity is supporting research on parallel computing necessary to ensure that the Department's programs requiring scientific supercomputing will have the tools to use efficiently the new machines as they become available and to exploit their maximum capability.

The DOE scientific research community comprises a large fraction of this nation's expertise in large scale computational modeling, with applications in fusion energy, weapons design, and fundamental physical sciences research. Much of the scientific research and development effort throughout DOE programs is focused directly on analytical and numerical modeling of physical processes. An understanding of the fundamental principles upon which these models are based is important for developing energy systems for the future. Thus, research in mathematical analysis, algorithms, and computational techniques forms the knowledge base for conducting scientific research. DOE's lead role in the development and application of supercomputing techniques is based on this fundamental research.

The subprogram funds basic research at many of the national laboratories, universities, and private research institutions in three major activities: Analytical and Numerical Methods, Information Analysis Techniques, and Advanced Computing Concepts. In addition, experimental computing capabilities were established as research facilities to support the exploration of new concepts in large scale scientific computing. Regular meetings with program managers in other Federal agencies help coordinate these activities.

In FY 1985, an enhancement in the research effort in large scale scientific computing was started, with emphasis on algorithms for large scale parallel architectures based on several experimental computer systems under study. The emphasis on new parallel multiprocessor architectures applies to all three activities of the research activity: analytical and numerical methods for the solution of systems of differential equations, information analysis techniques, and advanced computing concepts.

This activity will greatly accelerate the pace of projects in parallel supercomputing so that some of the research machines can be brought to a usable state as soon as possible. Also included in the long range plan for this activity is an expanded participation by interdisciplinary teams working on problem decomposition, algorithm development, numerical analysis, performance evaluation, and languages for highly parallel architectures.

The R&D projects supported by this activity have two important facets requiring substantial cooperation and coordination among traditionally separate groups. One facet is interdisciplinary teams of computational scientists, computer scientists, and mathematicians working on all aspects of large scale scientific computing problems. The other is the cooperation of industry, government and universities on the design and implementation of prototypes of several potentially strong candidate architectures; this will be an important proof of concept activity.

The university researchers play the major role in generating ideas and research software and in training graduate students in generating new applications. National laboratory staff are in the forefront of tackling real world, large scale scientific problems and have unique resources for participating in these research projects. Industry likewise has a unique role in providing state of the art production and testing facilities and stands to reap great benefits in understanding future architecture and software issues that tend to limit industry use of supercomputers currently. The transfer of technology from the academic and laboratory research environment to industry could be as rapid as possible through these cooperative projects. The DOE program will be coordinated with the other agencies to share common facilities where possible, such as very large scale integration (VLSI) design facilities and component fabrication.

Analytical and Numerical Methods

This activity supports the investigations of the fundamental properties of mathematical models used to simulate complex physical phenomena observed in energy systems. This category includes applied analysis, numerical methods for solving large systems of differential equations, and numerical analysis of algorithms.

These funds support interdisciplinary teams including analysts, applied mathematicians, computational scientists, postdoctoral fellowships and graduate students collaborating in universities and national laboratories to research parallel algorithms for mathematical software tools used in implementing these models. As documented in the David Committee's report, "Renewing U. S. Mathematics: Critical Resource for the Future," published by the National Research Council of the National Academy of Sciences, DOE's research program in applied mathematical sciences plays a lead role in support of basic research at the interface between mathematical and computational sciences. The shortage of graduate students and post-doctoral support in this crucial area is becoming severe and the current demand for trained computational scientists by industry is rapidly draining the academic pool and the pool of talent available to the National Laboratories. The intent is to strengthen the ties between universities and the Labs and to provide salary support and computational resources in the disciplines associated with computational science. The DOE program, with its unique history of basic research in these areas, is positioned to provide strong leadership in filling the "pipeline" from graduate studies to postgraduate studies to research positions in academia and the Labs.

Information Analysis Techniques

This activity supports the investigation of data management and statistical analysis of large scientific data sets common in DOE mission programs, such as High Energy Physics, Magnetic and Inertial Fusion, and others. Funding here supports advanced work in computational statistical methods and in techniques for managing the large scale scientific databases common to many DOE mission applications. The existence of a large set of valuable scientific databases in the DOE multiprogram and single purpose laboratories provides a unique resource for basic research in this area. Very few universities have the opportunity to study real scientific databases of the complexity found in our mission programs; a strong research effort in this area must be supported if huge data management and analysis problems are to be solved.

Advanced Computing Concepts

This activity meets a wide range of problems associated with the effective use of large scale scientific computers for the

Department's research and development program. The areas of research include software engineering techniques for design of quality mathematical software, distributed systems, and parallel architectures for high performance scientific computers. A variety of new techniques for decomposing large problems to execute on a large number of concurrent processors is being investigated in DOE laboratories and universities. These techniques will provide the insight needed to build the supercomputers of the 1990's.

These new architectures, employing perhaps hundreds or thousands of fast processors working concurrently to solve a single problem, require a thorough reexamination of all aspects of the solution method. Problems must be decomposed into parallel steps for concurrent execution. Algorithm characteristics such as accuracy, stability, robustness, and correctness of implementation must be investigated anew.

This activity supports joint university/industry/laboratory projects for the design phase of research parallel architectures started in FY 1985. Several projects have been supported over the past few years to design, analyze and simulate a variety of innovative multiprocessor architectures. Typically, these projects have been supported at the level of $100,000 to $500,000 per year to demonstrate the feasibility of implementing large scale computational models common in DOE programs. These projects are now expanding to the implementation phase to demonstrate and evaluate the capabilities of these architectures for future supercomputers. In FY 1985, several projects were supported at the level of between $1,000,000 and $2,500,000 per year to provide for implementation and software development of research machines. Each project represents a collaboration among universities, national laboratories, and private industry.

Summary of Applied Mathematical Sciences Research Projects
Fiscal Years 1984 and 1985

Institution	Project Manager/ Investigator(s)	Title of Project	Summary of Project in FY 1985	Start Date	End Date	FY 84 ($000)	FY 85 ($000)
** Budget and Reporting Activity: KC 070101							
Ames Laboratory, Iowa State University	James Corones/	Analytical and Numerical Methods - Wave Propagation	The Applied Mathematical Sciences program at Ames is exclusively in the area of applied analysis. This effort is focused on basic and applied problems of wave propagation. Particular attention is given to direct and inverse scattering theory and their applications. The purpose of this work is to develop mathematical procedures and tools that can be used to characterize and/or image nonuniformities in materials (density variations, stress fields, cracks, potential variations, for example). Two types of approaches are employed: 1) direct scattering, in which the nonuniformities are given and the objective is to compute the scattered field (e. g., of elastic waves scattered by cracks) and thereby determine characteristic signatures of scatterers; and 2) inversion procedures, in which incoming and outgoing signals are assumed known and the objective is to deduce the characteristics of the nonuniformities from this knowledge. Our inversion studies incorporate realistic constraints such as limited angle softview, limited apertures and noise. The techniques developed have wide application, including quantitative NDE (QNDE), geoprospecting, remote sensing of nuclear waste, plasma diagnostics, and fission reactor diagnostics. At present, QNDE is the major applications orientation of the effort. In addition, there is work on special functions directed at facilitating their use and computation.	10/01/85	09/30/86	400	475

Summary of Applied Mathematical Sciences Research Projects
Fiscal Years 1984 and 1985

Institution	Project Manager/ Investigator(s)	Title of Project	Summary of Project in FY 1985	Start Date	End Date	FY 84 ($000)	FY 85 ($000)
Argonne National Laboratory	Paul Messina/ Hans Kaper	Analytical and Numerical Methods	Applied analysis research involves the application of analytical and numerical techniques to problems in the natural and engineering sciences. Major efforts focus on spectral analysis of differential operators and bifurcation and stability analysis of nonlinear phenomena in combustion. In addition, research continues on modeling and analysis of certain fluid flow and materials science problems. Computational mathematics research involves the design and analysis of numerical algorithms, the development of special techniques to measure algorithm reliability and efficiency, and the preparation of software based on broadly portable algorithms for vector and parallel computers. New studies have also begun on defect migration phenomena, high dimensional quadrature, optimization techniques for Stirling engine design, and evaluation criteria for parallel algorithms. Use of our advanced computing research facility, as well as advanced computer systems at other sites, is essential for this investigation.	10/01/85	09/30/86	950	950
Atlanta University	Ronald Mickens/	Investigation of Perturbation Techniques for Nonlinear Difference Equations	Our major research activity will be the construction of a number of regular and singular perturbation techniques which may be used to obtain approximate analytical solutions to several classes of nonlinear difference equations. We will put particular emphasis on those procedures that lead to uniformly valid asymptotic solutions. An investigation will also be carried out on the modeling of nonlinear, first order ordinary differential equations by difference equations. In particular, procedures that give rise to models with zero local truncation error will be emphasized.	08/01/83	07/31/86	67	83

Summary of Applied Mathematical Sciences Research Projects
Fiscal Years 1984 and 1985

Institution	Project Manager/ Investigator(s)	Title of Project	Summary of Project in FY 1985	Start Date	End Date	FY 84 ($000)	FY 85 ($000)
Brookhaven National Laboratory	Ronald Peierls/ Robert Marr	Analytical and Numerical Methods	This task involves the application of analytical and numerical techniques to investigate and develop computational methods for problems that arise in a variety of scientific and engineering contexts of importance to DOE. The goal is to develop methods that are both reliable and efficient. Much of the current emphasis is on more realistic modeling of three dimensional problems and towards the understanding of nonlinear effects. Even the most efficient methods imply massive computational requirements which will only be met by computers with parallel architectures. Specific areas of research include the numerical modeling and solution of direct and inverse wave propagation problems; the discretization of linear and nonlinear elliptic boundary value problems using finite element and finite difference methods; and the application of iterative solution methods such as preconditioned conjugate gradient methods to the resulting systems of equations. In image reconstruction, development of algorithms for nuclear magnetic resonance imaging continues, with current emphasis being on the use of projection reconstruction methods for spatially imaging chemical shifts, the investigation of filtered back transform reconstruction algorithms and inherent performance characteristics of generalized Fourier methods, and the investigation of a new imaging strategy based on stochastic NMR techniques.	10/01/85	09/30/86	457	360

Summary of Applied Mathematical Sciences Research Projects
Fiscal Years 1984 and 1985

Institution	Project Manager/ Investigator(s)	Title of Project	Summary of Project in FY 1985	Start Date	End Date	FY 84 ($000)	FY 85 ($000)
California Institute of Technology	Herbert Keller/ Philip Saffman	Numerical Analysis, Computing and Fundamental Studies of Energy...	The work is concentrated in two main areas: hydrodynamics and numerical methods for the solution of partial differential equations. The hydrodynamics work includes vortex dynamics, turbulent flow, shear flow, and flow in porous media. The main emphasis on numerical methods development has been in the general area of path following and continuation methods, including global Newton and homotopy methods. These methods have been used along with finite difference and spectral methods for solving two point boundary value problems, partial differential equations and integral equations. Adaptive mesh selection and curvilinear overlapping grid methods are beginning to be used. Additional effort will be aimed toward the development of software for accessing remote large scale scientific computing resources.	04/01/83	03/31/86	259	271
Columbia University	Michael Tabor/	Singularity Structure of Nonlinear Ordinary and Partial Differential Equations	This project aims to study both nonlinear ordinary and partial differential equations from the point of view of their singularity structure. For ordinary differential equations the property that the solutions be meromorphic (in time) seems to provide a test of integrability for these systems. Our aim is to provide an in depth understanding of this technique and to extend its range of applicability. In the case of partial differential equations a similar concept has been developed involving the theory of functions of many complex variables. The validity of this technique, its extensions and its relationship to other methods in the field of integrable systems are to be investigated in detail. Overall, the program aims to provide a unified approach to a wide variety of nonlinear equations and the many physical processes that they model.	05/01/84	04/30/87	57	63

Summary of Applied Mathematical Sciences Research Projects
Fiscal Years 1984 and 1985

Institution	Project Manager/ Investigator(s)	Title of Project	Summary of Project in FY 1985	Start Date	End Date	FY 84 ($000)	FY 85 ($000)
Cornell University	Thomas Coleman/	Topics in Large Sparse Nonlinear Optimization	Many problems in the physical science, engineering, and econometric disciplines involve the numerical solution to a system of nonlinear equations or the computation of the minimizer of a nonlinear function. Often these nonlinear problems are of large dimension involving several hundred to several thousand variables. Typically, the relevant Jacobian and Hessian matrices are sparse. In addition, it is often the case that a single evaluation of the function or gradient is an expensive operation. The purpose of the proposed research is to explore ways in which sparsity can be exploited so that the number of function (or gradient) evaluations is 'minimized' and yet a solution is efficiently obtained. In particular, this project will examine several computational questions concerning the efficient estimation of a sparse symmetric Jacobian matrix by finite differences and the economic use of partial Jacobian information. Also, this research will examine several related issues that arise when minimizing a large sparse nonlinear function in the presence of constraint functions.	05/01/83	04/30/86	44	52
Cornell University	Mitch Feigenbaum/ Eric Siggia	Systems Near the Onset and in the Well Developed Regimes of Turbulence	We present a program of research dedicated to chaotic behavior in both the onset regime and in the fully developed (or high modal) regime. The investigations proposed entail a blend of a computational effort to gain data supporting the insights, and assisting the development, of a theoretical effort employing recently developed methodologies and ideas. The first aspect of this program - work pertaining to the onset of turbulence - is founded on the fact that universal minimal (and low) modal models can successfully embrace the behavior of real continuum systems. Our attack includes the investigation of chaos borne from low order quasiperiodicity; modifications of these phenomena in the presense of spatial symmetries; the continued investigation of functional integrals in order to determine the sensitivity of these processes to external noise and an attempt to develop descriptions of the fractional dimensional attractors that determine post onset turbulence. The second aspect of our research is centered on intrinsically high modal turbulence arising either from the onset behavior in large cells or from high Reynolds' number excitations. In the first case, we intend to continue investigations of theories to describe pre-Hopf bifurcation drifts in degenerate spatial structures. In the case of fully developed turbulence, we intend to perform investigations of details of the flow that can clarify the role of large scales in determining small scale intermittency.	01/01/83	12/31/85	171	187

Summary of Applied Mathematical Sciences Research Projects
Fiscal Years 1984 and 1985

Institution	Project Manager/ Investigator(s)	Title of Project	Summary of Project in FY 1985	Start Date	End Date	FY 84 ($000)	FY 85 ($000)
Cornell University	Geoffrey Ludford/	Special Year 1984-1985 on Reacting Flows: Combustion and Chemical Reactors	The Center for Applied Mathematics at Cornell will conduct a year long research meeting on reacting flows (combustion and chemical reactors). The Special Year will consist of eight distinguished speakers brought in during the academic year, culminating in the two week AMS-SIAM Summer Seminar, which was awarded to Cornell for this purpose. The object is to synthesize the newly emerged theory of reacting flows up to its interface with the computation of such flows and to preserve the knowledge gained from this activity in a Proceedings volume that will provide a firm foundation for future research in the computational aspects of reacting flows. (Note: This Special Year was jointly funded by Applied Mathematical Sciences, Engineering Sciences, and Chemical Sciences.)	09/01/84	08/31/85	23	0
Indiana University	Cyprian Foias/	Criteria for the Reliability of Numerical Approximations to Fluid Flow Problem	The numerical approximation of the solutions of fluid flow models is a difficult problem encountered in many areas of energy research. In all numerical methods implementable on digital computers, the basic question is how to choose the number of elements, N say (Galerkin modes, finite difference cells, finite elements, etc.) in order to guarantee that the approximation at least behaves like the exact solution. In the absence of a rational prescription for the determination of this number until now, it has been estimated either empirically or on the basis of physical intuition. The purpose of this research is to derive useful estimates for N from some newly discovered fundamental properties of the viscous incompressible Navier Stokes equations.	05/01/82	08/31/85	72	0

Summary of Applied Mathematical Sciences Research Projects
Fiscal Years 1984 and 1985

Institution	Project Manager/ Investigator(s)	Title of Project	Summary of Project in FY 1985	Start Date	End Date	FY 84 ($000)	FY 85 ($000)
Kent State University	Richard Varga/ Paul Wang	Applications of Numerical Analysis and Computer Symbolic Computation	The emphasis of this research is on the theoretical basis for workable numerical algorithms for general iterative methods in linear algebra and for coupling of the construction of these algorithms with symbolic algebra systems. Three areas are under investigation: (1) An asymptotic analysis of basic iterative methods for solving matrix equations arising primarily from the use of finite element methods; (2) The use of symbolic and numerical techniques to attempt to settle longstanding conjectures in function theory and in approximation theory; and (3) The integration of symbolic and numerical computing techniques.	12/01/83	11/30/86	103	108
Lawrence Berkeley Laboratory	Paul Concus/ Alexandre Chorin	Analytical and Numerical Methods	The analytical and numerical methods program at LBL focuses on the development of techniques for use in mathematics, physics, and engineering. Common to the various projects is the design of new methods that utilize computer resources effectively to provide fast and efficient algorithms, with the goal of illuminating increasingly complex and realistic physical and engineering phenomena. The physical topics include fluid dynamics, high speed gas flow, turbulence theory, the dynamics of fronts and interfaces, combustion, capillarity, and image reconstruction. The techniques developed and used include vortex methods, high resolution methods for hyperbolic equations, fast enumeration methods, numerical integration in function spaces, and iterative methods for algebraic systems.	10/01/85	09/30/86	645	600

Summary of Applied Mathematical Sciences Research Projects
Fiscal Years 1984 and 1985

Institution	Project Manager/ Investigator(s)	Title of Project	Summary of Project in FY 1985	Start Date	End Date	FY 84 ($000)	FY 85 ($000)
Lawrence Livermore National Laboratory	Robert Huddleston/ James McGraw	Computational Mathematics	Computational mathematics research focuses on the design, development, and testing of algorithms to solve problems more accurately and/or more efficiently than existing methods, and to solve problems that have been intractable using existing algorithms. Research on adaptive methods for solving PDEs on multiprocessors will be expanded to include local grid refinement methods. Research on adaptive methods involving moving grids will continue. In addition, research will continue on the multiprocessor solution of algebraic systems arising from an implicit time step on a large stiff system of ODEs, and on shape preserving and scattered data interpolation and approximation methods for bivariate functions. In all of these areas attention will be focused on algorithm development and implementation on multiprocessors.	10/01/85	09/30/86	175	250
Los Alamos National Laboratory	George Bell/ Basil Nichols	Numerical Algorithms and Nonlinear Mathematics of Energy Systems	In adaptive numerical methods, we are constructing both static and dynamic rezoning algorithms; we are addressing the concept of dynamic processor allocation strategy for their implementation on machines with advanced computer architectures. Other developments include: composite grids; interface tracking through adjustment of the new grid on the moving boundary; high order collocation methods and supra-convergent difference schemes on irregular grids; semi-coarsening in multigrid methods and blackbox multigrid software. We are concurrently developing and demonstrating the concept of adaptive mesh refinement using the multigrid method; our testbed is an existing highly vectorized 3-D Eulerian multigrid finite element code. We are investigating 3-D front tracking methods as to their compatibility with the multigrid framework. We expect the implementation of the multigrid framework to be highly effective on machines with massively parallel architectures. Jointly with the Center for Nonlinear Studies at Los Alamos, the main thrust of our effort in applied analysis is directed at investigating pattern selections, instabilities, and chaos on strongly dissipative interfaces in flows. For this class, we establish genuine low modal behavior and we construct explicit formulas for the low fractal dimension of the strange attractor as a function of a typical pattern size, in the post onset turbulence regime. We demonstrate that these PDE´s are rigorously equivalent to a low dimensional dynamical system; we establish the existence of a universally attracting, compact, finite dimensional inertial manifold. Future efforts are aimed at further bridging the gap between dynamical systems and PDE´s.	10/01/85	09/30/86	325	350

Summary of Applied Mathematical Sciences Research Projects
Fiscal Years 1984 and 1985

Institution	Project Manager/ Investigator(s)	Title of Project	Summary of Project in FY 1985	Start Date	End Date	FY 84 ($000)	FY 85 ($000)
New York University, Courant Institute	Peter D. Lax/ James Glimm	Analytical and Numerical Methods	Two remarkable developments have taken place in many of the high technologies: the replacement of much experimentation by computer simulation, and the digital storing and processing of information. The reason for the former is the enormous gain in economy and flexibility of computer simulation over experimentation. This is particularly true in the energy related technologies; it is worth pointing out that the AEC was among the first of the technologically oriented organizations to recognize that in order to carry out its mission, it had to rely on a tremendous amount of computer modeling. A very large percentage of the relevant models are described in terms of partial differential equations. The success of solving these equations depends critically on the availability of new computing methods and efficient algorithms for implementing them. This is borne out by a glance at the history of scientific computing; it shows that each advance was at least partly fueled by an improved numerical method. As examples of such ideas, one can point to shock capturing, alternating direction, high order difference methods, finite elements, use of complex coordinates, random choice, discrete vorticity, spectral methods, implicit methods, flux corrected transport, adapted mesh, and others. In the last five years this Laboratory has been active in the design of a novel highly parallel computer. In the next two years we expect to devote a significant portion of our effort to devising algorithms suitable for such an architecture. The Courant Institute was always deeply involved in the invention and development of methods for solving partial differential equation, and this pursuit continues to be a significant portion of the scientific program of the Courant Mathematics and Computing Laboratory.	10/01/85	09/30/86	1210	1162

Summary of Applied Mathematical Sciences Research Projects
Fiscal Years 1984 and 1985

Institution	Project Manager/ Investigator(s)	Title of Project	Summary of Project in FY 1985	Start Date	End Date	FY 84 ($000)	FY 85 ($000)
Northwestern University	Bernard Matkowsky/	Research in Nonlinear Problems of Energy	Faculty associates on this project are Profs. M. Matalon, T. Erneux, Z. Schuss, and A. van Harten. A concerted attack on a large variety of problems involving the transition from laminar to turbulent combustion will be made. The program will involve several faculty members and research associates including postdoctoral students. The program is a theoretical effort, closely coordinated with experimental work. In this program, several relevant problems that are governed by systems of nonlinear ordinary and partial differential equations will be studied. Specific problems under investigation are: application of the Matalon-Matkowsky model of flames with nonzero thickness to the stability of curved flames; secondary bifurcation from cellular or pulsating flames; downward flame propagation in long vertical channels; burner stabilized flames in more than one dimension; interaction of flames with sound; and application of stochastic differential equations to the problem of relative stability of various multiple stable states of a system, including the derivation of more accurate expressions for chemical reaction rates.	01/01/84	12/31/87	100	105

Summary of Applied Mathematical Sciences Research Projects
Fiscal Years 1984 and 1985

Institution	Project Manager/ Investigator(s)	Title of Project	Summary of Project in FY 1985	Start Date	End Date	FY 84 ($000)	FY 85 ($000)
Oak Ridge National Laboratory	Robert Ward/ Michael Heath	Computational Mathematics	Research in computational mathematics is concentrated in numerical linear algebra, with the focus on problems involving sparse matrices in the areas of systems of linear equations, eigenvalue problems, and least squares approximations. Our current emphasis is on development of both general methodology and specific algorithms for matrix computations on parallel computers. In two separate, but related, lines of research we are developing both data flow and control flow algorithms for parallel computation of the Cholesky factorization. This work will be extended to include both symbolic and numeric phases of sparse factorization. We are also developing a block Jacobi algorithm for symmetric eigenvalue problems for parallel computation using systolic arrays. Our work on the M matrices arising from queuing networks and Markov chains will continue, including the analysis of both direct and iterative methods for solving homogeneous linear systems involving such matrices, as well as updating sensitivity analysis, and iterative refinement.	10/01/85	09/30/86	210	290
Oak Ridge National Laboratory	Robert Ward/ Allan Solomon	Applied Analysis	Research in applied analysis is carried out in two problem areas: moving boundary problems and stochastic analysis, with primary emphasis on the first topic. Work in moving boundary problems concerns the analysis of heat and mass transfer processes involving a change of phase. The mathematical questions are those associated with properly posing and analyzing, both qualitatively and quantitatively, the partial differential equations used to represent the idealized process. Our attention has been focused on modeling binary alloy solidification, with emphasis on the study of "mushy" zones and the initial behavior of phase fronts. This work will be complemented with new studies, such as hyperbolic heat transfer. Work in stochastic analysis concerns the modeling of disordered systems such as liquids and amorphous solids. Here the mathematical questions are associated with the spectral properties of certain operators which give information from the mathematical models of specific systems. This work has also become involved with alloy solidification, but with nonequilibrium or "ultrafast" solidification resulting in amorphous alloys.	10/01/85	09/30/86	205	210

Summary of Applied Mathematical Sciences Research Projects
Fiscal Years 1984 and 1985

Institution	Project Manager/ Investigator(s)	Title of Project	Summary of Project in FY 1985	Start Date	End Date	FY 84 ($000)	FY 85 ($000)
Rice University	Richard Tapia/ John Dennis	Quasi-Newton Methods for Unconstrained and Constrained Optimization	Researchers associated with this contract include Trond Steihaug of Rice University and Homer Walker of the University of Houston and their graduate students. The objective of this work is to extend the theory of quasi-Newton methods for unconstrained and constrained optimization. In the area of unconstrained optimization we have produced some theory that covers the situation when the objective function is known to possess noise and cannot be accurately evaluated. In constrained optimization we have developed a general convergence theory and have derived several new algorithms.	08/01/82	12/31/85	0	84
Sandia National Laboratories, Albuquerque	Lawrence Shampine/	Mathematical Software	Fundamental research in the numerical solution of ODEs leading to improved understanding of the task and to more efficient and reliable mathematical software: New understanding and improved estimates of local errors will be achieved. Several kinds of standard methods will be made more efficient and reliable. Global (true) error estimates will be devised. New methods for the solution of stiff problems will be proposed. A comprehensive package of Bessel functions of complex argument and nonnegative orders will be constructed. For maximum applicability and efficiency, each code requires the use of many methods. This necessitates fundamental mathematical and software research as well as extremely careful development and testing. Mathematical tools for solving inverse problems involving systems of ODEs will be developed. These tools will include better methods for solving constrained optimization problems and more efficient methods for solving systems of ODEs. Resulting codes will be used in the development of software for solving chemical kinetics parameter estimation problems and seismic parameter estimation problems.	10/01/85	09/30/85	210	225

Summary of Applied Mathematical Sciences Research Projects
Fiscal Years 1984 and 1985

Institution	Project Manager/ Investigator(s)	Title of Project	Summary of Project in FY 1985	Start Date	End Date	FY 84 ($000)	FY 85 ($000)
Sandia National Laboratories, Livermore	Robert Kee/ Linda Petzold	Analytical and Numerical Methods	This research is concerned with the solution of disparate scale partial differential equations; the approach is both analytical and computational. Although the results of this work are applied generally, one unifying theme is the interest in the problems that arise in combustion modeling. On the analytical side we are developing asymptotic models that explain instabilities in flames, alloying reactions, and solid propellants. On the computational side, we concentrate on large systems of stiff PDE´s, particularly those that occur in models involving systems of complex chemical reactions. To achieve accuracy and efficiency in time dependent problems we are developing implicit methods, such as backwards differentiation formulas and implicit Runge Kutta, that deal effectively with the widely varying time scales that are present in chemical kinetics models. To achieve spatial accuracy and computational efficiency, we are continuing to develop adaptive gridding methods to resolve any spatially localized behavior, such as a flame. An important part of our program is the application of our work to practical problems, especially in chemical sciences and combustion.	10/01/85	09/30/86	435	450

Summary of Applied Mathematical Sciences Research Projects
Fiscal Years 1984 and 1985

Institution	Project Manager/ Investigator(s)	Title of Project	Summary of Project in FY 1985	Start Date	End Date	FY 84 ($000)	FY 85 ($000)
Science Applications International Corp	Alan Drobot/ Robert Gelinas	Moving Finite Elements in 2-D	This research project on moving finite element methods is focused on research and development of the method in two dimensions, including the development of research oriented software for solving systems of partial differential equations describing physical problems with shocks and moving fronts, such as in combustion processes. The project leadership was assumed by Drobot in 1985 with the purpose of concentrating on schemes for adding, removing, and reconnecting nodal points and gaining some understanding of the consequences on the physical accuracy of the solutions obtained. They offer to provide a listing of the moving finite element code as well as a users' manual. A manuscript on the properties of this method in 2D using a triangular grid should be available by the end of the contract.	04/15/81	07/31/85	65	36
Stanford University	George Dantzig/	Systems Optimization Project	Co-investigators on this project are R. W. Cottle, Walter Murray, and B. Curtis Eaves. Research associates are Philip Gill, Michael Saunders, Margaret Wright, and John Stone. Research topics include time staged linear programs, analysis of standard and ellipsoidal linear programming algorithms, methods for large scale linear complementarity problems, unconstrained optimization, constrained optimization, automatic scaling techniques, and simplified algorithms for solving dynamic multiregional planning models. The overall objective of the research is to develop effective methods for solving optimization (mathematical programming) problems. The areas of special concentration include staircase (dynamic) linear programs, large scale quadratic programming, nonlinearly constrained optimization (both dense and large scale), large scale energy-economic models, linear complementarity problems, and equilibrium problems.	11/01/83	10/31/86	300	315

Summary of Applied Mathematical Sciences Research Projects
Fiscal Years 1984 and 1985

Institution	Project Manager/ Investigator(s)	Title of Project	Summary of Project in FY 1985	Start Date	End Date	FY 84 ($000)	FY 85 ($000)
University of California, Berkeley	Jerrold Marsden/ Alan Weinstein	Stability of Fluids, Plasmas, and Nonlinear Wave Equations	The energy-Casimir method developed by the proposers, D. Holm, T. Ratiu and others will continue to be applied to a number of concrete fluid and plasma problems. Using these methods, free boundary problems will be investigated, as well as stratified flow. Related work will be carried out on averaging techniques and limits of projections of Hamiltonian structures, such as occur in BBGKY hierarchy to Maxwell-Vlasov and the Maxwell-Vlasov to charged fluid transitions. In addition, the proposers plan to use their techniques to study nonlinear wave equations in conjunction with the first year's planned visitor. Questions of stability, periodic solutions and chaos, (the latter using the Holmes-Marsden work on Melnikov's methods), will be considered. Melnikov's method is also expected to be useful in studying chaos for forced free boundary problems.	07/01/85	06/30/88	97	145
University of Illinois	C. W. Gear/ Robert Skeel	Computational Methods and Software for Ordinary Differential Equations	This work involves research on several open questions in the area of ODEs, viz., how to handle problems with multiple periods, why constant stepsize methods seem to work for certain classes of differential algebraic equations (DAEs), and how to handle general DAEs without prior symbolic analysis. Also, application of iterative methods to the solution of the nonlinear equations arising at each step in a stiff ODE or DAE and the effect of discontinuities on automatic stepsize methods will be studied. This rather broad scope covers the entire spectrum of problem areas encountered in solving energy related problems described by ODEs (and some classes of PDEs solvable by the Method of Lines).	11/16/83	11/15/86	184	194

Summary of Applied Mathematical Sciences Research Projects
Fiscal Years 1984 and 1985

Institution	Project Manager/ Investigator(s)	Title of Project	Summary of Project in FY 1985	Start Date	End Date	FY 84 ($000)	FY 85 ($000)
University of Maryland	James Yorke/ Edward Ott	Study of Chaotic Dynamics and Dimension of Strange Attractors	Faculty associate on this contract is Professor Sheldon Newhouse. Celso Grebogi is a research associate. Even simple mathematical models of physical systems are often observed to exhibit rather complex time evolution. In such situations the time evolution is often labeled "chaotic" or "turbulent." In many cases this chaotic behavior is associated with rather intricate, arbitrarily fine scaled, geometric structure of the attracting set to which a trajectory tends as the system evolves. These sets often possess Cantor set structure and are called "strange attractors." The study of strange attractors has recently undergone explosive growth. Motivation for this comes partly from the fact that strange attractors are being found to be of fundamental importance in many branches of science and engineering. Examples illustrating the wide ranging applications of strange attractors to physical and engineering problems are studies of turbulence in fluids, electrical circuits with tunnel diodes, buckling beams, plasma physics, lasers, solid state physics, magnetohydrodynamics, oscillations in the earth's magnetic field, chemically reacting systems, and many others. This program is to study two of the important aspects of strange attractors: the dimension of the strange attractors and how strange attractors evolve as a parameter is changed. Numerical experiments are a major tool in this study.	05/01/83	04/30/86	138	155

Summary of Applied Mathematical Sciences Research Projects
Fiscal Years 1984 and 1985

Institution	Project Manager/ Investigator(s)	Title of Project	Summary of Project in FY 1985	Start Date	End Date	FY 84 ($000)	FY 85 ($000)
University of Texas	David Young/ Graham Carey	Iterative Solution and Finite Element Analysis in Computational Mechanics	Co-investigators on this project are David Kincaid and Kamy Sepehrnoori. This project represents a research program embracing interrelated topics in applied mechanics, numerical analysis, and computer science. Specifically, we are investigating the theory and use of iterative solution algorithms in conjunction with finite element methods, for computing numerical solutions of certain classes of transient and nonlinear problems in mechanics and heat transfer that are of significant research and practical importance. Concurrently, we shall address related questions of accuracy and error estimates for finite element approximation and of stiffness and stability for semidiscrete systems in the transient problems. The actual solution of nonlinear mechanics problems, such as those encountered in Navier-Stokes problems, in reservoir simulation and in heat transfer with nonlinear chemical reactions will provide an assessment of new methodology arising from the research and for realistic applications. Finite element and iterative solution algorithms will be developed and evaluated for computations on supercomputers, vector processors as well as other advanced architectures. Part of the work involves a major expansion of the ITPACK library, with emphasis on improved vectorization and the use of special finite element methods adapted to vector computers. The suitability of certain types of iterative methods in conjunction with finite element methods for state of the art emerging parallel computing configurations is also to be studied.	08/01/84	07/31/87	135	155

Summary of Applied Mathematical Sciences Research Projects
Fiscal Years 1984 and 1985

Institution	Project Manager/ Investigator(s)	Title of Project	Summary of Project in FY 1985	Start Date	End Date	FY 84 ($000)	FY 85 ($000)
University of Washington	Jirair Kevorkian/	Resonance in Weakly Nonlinear Systems	Resonance is a common phenomenon in any physical process that involves coupled oscillations. In this research we consider recent work on nonlinear resonance in systems with slowly varying frequencies and indicate how this work applies to a quantitative description of the operation and stability of free electron lasers. Related to this is work in deriving adiabatic invariants for systems passing through resonance and work on wave propagation problems using multiple variable expansion techniques. In this project we will investigate the following areas: (1) A systematic analysis of free electron lasers including higher order effects and stability; (2) A study of nonlinear evolution equations as they occur in (1) or more generally in systems with strong interactions between modes; (3) The influence of nonlinear dissipation; and (4) Resonant wave interactions in unbounded media with slowly varying properties.	08/01/83	07/31/86	40	45
Virginia Polytechnic Institute & State University	Paul Zweifel/ William Greenberg	Applied Mathematics of Transport Theory	The goals of this research are to develop applicable mathematical techniques that are relevant to the physical transport processes underlying energy production. These include, for example, neutron transport, thermonuclear plasmas, rarefied gas dynamics, etc. Our concentration is on phonon and electron transport in solids, evaporation and condensation, study of abstract kinetic equations (existence and uniqueness theory), nonlinear plasma stability and acceleration and homogenization techniques in numerical transport theory calculations. Both analytical, function theoretical, techniques and numerical computations are planned.	02/01/83	11/30/86	51	92

Summary of Applied Mathematical Sciences Research Projects
Fiscal Years 1984 and 1985

Institution	Project Manager/ Investigator(s)	Title of Project	Summary of Project in FY 1985	Start Date	End Date	FY 84 ($000)	FY 85 ($000)
Washington State University	Kuruppu Ariyawansa/	Conic Algorithms for Unconstrained Minimization - Large Sparse Hessians	Existing algorithms for unconstrained minimization are based on local quadratic approximations. Local quadratic approximants cannot incorporate information on objective function values. Methods like DFP and BFGS that are successful for small dense problems cannot be fully extended to incorporate information on the structure of the Hessian in large sparse problems. These shortcomings are due to the limited degrees of freedom available in the quadratic model. The investigator has demonstrated that local conic approximations can be used to incorporate information on objective function values and to have the resulting algorithms possess good convergence properties. The objectives of the present research are three-fold. First we propose to continue the theoretical study of conic algorithms mentioned above to settle several issues necessary to implement these algorithms. Second, we propose to analyze a class of conic algorithms that includes complete conic extensions of DFP and BFGS methods for problems with sparse Hessians. Third, we propose to implement and thoroughly test several promising conic algorithms.	08/01/85	07/31/87	0	78

Summary of Applied Mathematical Sciences Research Projects
Fiscal Years 1984 and 1985

Institution	Project Manager/ Investigator(s)	Title of Project	Summary of Project in FY 1985	Start Date	End Date	FY 84 ($000)	FY 85 ($000)
Yale University	Tony Chan/	Efficient Numerical Methods for Large Nonlinear Systems	This project will study efficient numerical algorithms for solving large nonlinear systems that arise frequently in large scale scientific computing. In addition to general large nonlinear systems, we also consider problems with more inherent structures, including parameterized nonlinear systems that arise in homotopy and continuation methods and coupled nonlinear partial differential equations. The numerical techniques include multigrid and continuation methods, an approximate Newton method recently developed by the PI for solving coupled nonlinear systems and a nonlinearly preconditioned conjugate gradient method that does not require explicitly computing or storing the Jacobian. We will investigate the application of these methods to numerical continuation methods applied to the Navier-Stokes equations, the transonic flow equations and structural mechanics problems and to solving the coupled partial differential equations that arise in semiconductor device simulations. Finally, we will carry out the numerical implementations on the ELI CIRCUS, a ring of array processors being developed at Yale, and try to identify algorithmic implementations that can take advantage of the parallelism and vectorization capabilities provided by the architecture.	01/01/85	12/31/87	0	113

Summary of Applied Mathematical Sciences Research Projects
Fiscal Years 1984 and 1985

Institution	Project Manager/ Investigator(s)	Title of Project	Summary of Project in FY 1985	Start Date	End Date	FY 84 ($000)	FY 85 ($000)
** Subtotal **						7128	7603
** Budget and Reporting Activity: KC 070102							
Brookhaven National Laboratory	Ronald Peierls/ Herbert Robbins	Statistical Methods	The primary emphasis of this research is on the development of innovative methodologies for analysis of data sets that are restricted, incomplete, or complex in other ways. A characteristic of such problems is that the observations to be analyzed are on subjects that are valuable, irreplaceable, and not easily reproducible. Such characteristics apply to weapons tests, nuclear reactors, large scale engineering structures and other systems of importance in energy and environmental research. To obtain reliable results from the analysis of the limited data sets associated with such exotic systems requires more sophisticated analysis than more conventional statistics, and is likely to call for extensive computation, both to implement these more sophisticated methods and also in their development through simulation and Monte Carlo methods.	10/01/85	09/30/86	343	275
Brookhaven National Laboratory	Ronald Peierls/ H. Bruce Stewart	Display and Analysis Systems	This task uses numerical simulation and computer visualization to understand qualitative behavior of nonlinear dynamical systems. The goal is to develop and implement new visual approaches to understanding topological structures in the phase space of a dynamical system. The methods are applied to differential equation models from structural and fluid mechanics, electromagnetics, chemical kinetics, and biological and ecological oscillations or fluctuations. Application of these phase space geometric analysis methods yields information about underlying order in apparently random fluctuating behavior (chaotic attractors) and changes of stability in dynamic behavior (bifurcations). Phase space geometric analysis is vastly enhanced by the use of interactive computer graphics. Current effort is directed toward designing software that tightly integrates the functions of dynamic simulation (solving differential equations), constructing systematic phase space portraits, and viewing and editing the resulting curves and surfaces.	10/01/85	09/30/86	0	25

Summary of Applied Mathematical Sciences Research Projects
Fiscal Years 1984 and 1985

Institution	Project Manager/ Investigator(s)	Title of Project	Summary of Project in FY 1985	Start Date	End Date	FY 84 ($000)	FY 85 ($000)
George Washington University	James Foley/	Scientific Work Station Research	Scientific work stations, which are powerful networked personal computers with high quality graphics interfaces, are becoming an important tool for scientists to use in their daily computational, information management, and communications work. This research centers on several issues where progress will help increase the productivity, usability, programmability, and efficiency of these work stations. The research topics are: development and testing of a graphics interface to the work station and network operating system to augment current text oriented interfaces; development of a "look up table compiler" language as part of our work on device independent graphics for raster displays; and development of an object oriented graphics programming language and language implementation suited for use in the next generation of multiprocessor scientific work stations.	08/01/83	07/31/86	55	60
Lawrence Berkeley Laboratory	Carl Quong/ Arie Shoshani	Large Scale Scientific Data Management Research	The data management research program is aimed at the development of specialized data management techniques for scientific and statistical databases. Scientific databases result from experiments and simulations of physical phenomena, such as the collision of particle beams or modeling fluid dynamics systems. Statistical databases are collected primarily for the purpose of providing statistical summaries, such as energy consumption and production. These databases have characteristics that cannot be supported efficiently by existing commercial data management systems. The purpose of the data management research program is twofold: to identify the special characteristics and problems that exist in scientific and statistical databases, and to develop techniques that are tailored to the special structures and usage of such databases. Special emphasis is placed on techniques that can take advantage of parallel processing opportunities.	10/01/85	09/30/86	500	525

Summary of Applied Mathematical Sciences Research Projects
Fiscal Years 1984 and 1985

Institution	Project Manager/ Investigator(s)	Title of Project	Summary of Project in FY 1985	Start Date	End Date	FY 84 ($000)	FY 85 ($000)
Oak Ridge National Laboratory	Robert Ward/ Toby Mitchell	Statistical Methods	The statistical methods research effort is divided between the computational statistics (CS) and analysis of large data sets (ALDS) research projects. Research in the CS project will include work in computer aided experimental design, algorithms for computing distributional properties of estimators and test statistics, iteratively reweighted least squares algorithms, computational aspects of large scale regression problems requiring sparse matrix technology, and algorithms for computing Bayesian posterior distributions. Since most of these algorithms are computationally intensive, special emphasis will be placed on their implementation using parallel processing. Research in the ALDS project will focus on the following areas: evaluation of the effect of data transformation on misclassification probabilities in multivariate discriminant analysis; investigation of Bayesian variable selection methods along with the development of the necessary computer algorithms for computation of posterior distributions of interest; and development of methods to analyze data arising from imperfect sampling frames.	10/01/85	09/30/86	365	360

Summary of Applied Mathematical Sciences Research Projects
Fiscal Years 1984 and 1985

Institution	Project Manager/ Investigator(s)	Title of Project	Summary of Project in FY 1985	Start Date	End Date	FY 84 ($000)	FY 85 ($000)
Pacific Northwest Laboratory/ Battelle	Wesley Nicholson/ Marvin Erickson	Analysis of Large Data Sets	Effectively utilizing the growing quantities of complex energy related data requires new techniques in data analysis, data display, and data management. This integrated statistics and computer science research project is supported by the research computing facility, a statistical laboratory that serves as a testbed for new methods of and approaches to analysis. Specific topics addressed in graphical analysis research are 1) new techniques for increasing the number of points and dimensions that can be displayed and manipulated, and 2) the definition and automatic detection of "interesting" structure. Data analysis management research focuses on the management and control of 1) information generated by the analysis process (derived data) and 2) information concerning the analyst's interpretation of results. The specific issues that we will address are automated structure detection in large multivariate data sets and the applicability of supercomputers for these calculations.	10/01/85	09/30/86	500	400
University of Florida	Stanley Su/	A Dynamic Multicomputer System for Managing Scientific and Engineering Data	This research project entails the design, analysis, and evaluation of a multicomputer system that consists of multiple processors on a dynamically partitionable network with switchable main memory modules. They will investigate the architectural features of this system suitable for parallel processing of numerical and symbolic data for scientific and engineering databases common in energy research and applications. The design and evaluation of software algorithms that take advantage of these architectural features will be undertaken.	08/01/84	07/31/86	81	85

Summary of Applied Mathematical Sciences Research Projects
Fiscal Years 1984 and 1985

Institution	Project Manager/ Investigator(s)	Title of Project	Summary of Project in FY 1985	Start Date	End Date	FY 84 ($000)	FY 85 ($000)
University of Utah	Robert Barnhill/	Computer Aided Surface Representation	Faculty associates on the project are P. Alfeld, G. Farin, and A. Worsey. The aims of this research project are the creation of new surface forms and the determination of geometric and physical properties of surfaces. The investigation of new surface forms include the topics of piecewise polynomial triangular interpolants, monotone and convex surfaces, closed and convolution surfaces, tetrahedral interpolants, surfaces defined on surfaces, and a multidimensional Coons' patch (Gregory's Cube). The geometric and physical properties considered include the estimation of derivative data, numerical integration, contouring, and the tesselation of points in 3D. The surfaces to be created must satisfy a variety of criteria such as sufficient smoothness, reasonable shape, and sufficient accuracy or fidelity. We take the approach of developing a robust class of surfaces that can utilize arbitrarily located information that is frequently demanded in DOE applications. This approach is particularly difficult for multidimensional interpolation to data in three or more variables. The parallel rendering of surfaces will be investigated.	06/01/85	05/31/88	132	307

Summary of Applied Mathematical Sciences Research Projects
Fiscal Years 1984 and 1985

Institution	Project Manager/ Investigator(s)	Title of Project	Summary of Project in FY 1985	Start Date	End Date	FY 84 ($000)	FY 85 ($000)
University of Washington	Werner Stuetzle/ Andreas Buja	Nonparametric Methods in Multivariate Analysis	The principal goal of this research is the creation of new statistical tools to explore the structure of large scientific data sets. An initial focus of this work will be the detection of nonlinear relations amongst predictor and response data, where the response is a separable or additive function of the predictor variables. The problems to be studied involve the optimal determination of the separable response functions, a computationally intensive problem. A second area is the determination of whether the predictions obtained from a given data analysis are degenerate. This work generalizes the search for linear dependencies in classical regression analyses. In addition to the development of analytical models, the investigators will develop sophisticated graphics capabilities to explore the structure of large scientific data sets and assess model predictions generated computationally.	07/01/85	06/30/88	0	250
University of Wisconsin	David DeWitt/ Douglas Bates	Database Machines for Large Statistical Databases	This project will investigate the issues involved in the design of a high performance statistical database machine. One task is to perform a static and dynamic analysis of the principal statistical operations to determine the frequency of use of primitive linear algebra building blocks, such as matrix multiplication and inversion. The second task is the implementation of a relational dataflow database machine, called FLASH, on the reconfigurable multiprocessor hardware called CRYSTAL. FLASH will be modified to incorporate the linear algebra building blocks as new query processor primitive operations. The resulting system will be monitored as it is used by staff and students in the University of Wisconsin Statlab for the purpose of gathering statistics on various system performance parameters such as disk traffic, communications traffic among the various components, and utilization and consumption of resources by the linear algebra building blocks selected. A benchmark relational database will be developed and a set of benchmark queries will be designed to quantify performance features of various algorithms on various architectures in the context of a large statistical database.	01/16/83	01/15/86	86	121

Summary of Applied Mathematical Sciences Research Projects
Fiscal Years 1984 and 1985

Institution	Project Manager/ Investigator(s)	Title of Project	Summary of Project in FY 1985	Start Date	End Date	FY 84 ($000)	FY 85 ($000)
** Subtotal **						2062	2408
** Budget and Reporting Activity: KC 070103							
Argonne National Laboratory	Paul Messina/	Advanced Computer Systems Concepts	Research in automated reasoning plays a major role in our software engineering activities. Studies include development of an abstract machine for logic programming, design of new inference rules and strategies, and applications to problems of interest to DOE. A second major focus is advanced scientific computing. This program comprises the establishment of an advanced computer research facility and the development of a broad research program in mathematics and computer science. The goal is to understand and express the level of parallelism possible in numerical and nonnumerical tasks and to exploit the strengths of particular architectures. In addition, Argonne is working on applications of software engineering techniques to the production of program packages like MINPACK; abstract programming; and standards for Fortran, PL/I, and IEEE floating point arithmetic. New work includes design of simplification transformations for our automated program transformation system and creation of a unified package for eigenvalues and systems of linear equations.	10/01/85	09/30/86	1085	1950

Summary of Applied Mathematical Sciences Research Projects
Fiscal Years 1984 and 1985

Institution	Project Manager/ Investigator(s)	Title of Project	Summary of Project in FY 1985	Start Date	End Date	FY 84 ($000)	FY 85 ($000)
California Institute of Technology	Geoffrey Fox/	The Caltech Concurrent Computation Program	Associated faculty on this project are Aron Kupperman, Rob Clayton, Herbert Keller, Philip Saffman, R. Blandford, N. Corngold, R. Gould, P. Haff, B. Hager, P. Liewer, and T. Tombrello. The Caltech Concurrent Computation Project (CCCP) is a new approach to homogeneous architectures begun in the Caltech Cosmic Cube Project. An important lesson learned from that project is the importance of having a significant amount of computational power in each of the nodes (including a large enough memory local to the node to contain a significant computational problem) and the effectiveness of offloading the communication processes to coprocessors, including broadcast capabilities. The new nodes proposed, built from off the shelf components available from the vendors, will have these important features. In particular, the independent communication coprocessors should be capable of demonstrating the scalability of the hypercube architecture to much larger machines and also the feasibility of using heterogeneous processors in a hypercube configuration. They should be able to demonstrate true MIMD capability with this new processor node. The broad spectrum of scientific disciplines represented in this project is extraordinary (possibly unique). They have been successful in demonstrating techniques for mapping a long series of computational methods onto the hypercube architecture. Given the importance to DOE research and development programs of future parallel architectures for supercomputers, the Caltech project has pioneered a novel method for analyzing the structure of scientific models and devising new (and revising old) parallel algorithms to maximize the efficient use of these new machines.	08/01/85	07/31/90	900	1400

Summary of Applied Mathematical Sciences Research Projects
Fiscal Years 1984 and 1985

Institution	Project Manager/ Investigator(s)	Title of Project	Summary of Project in FY 1985	Start Date	End Date	FY 84 ($000)	FY 85 ($000)
Lawrence Berkeley Laboratory	Carl Quong/ Joseph Sventek	Advanced Computer Concepts	The advanced computer concepts program concentrates on the understanding and resolution of problems that prevent the effective use of distributed computer systems. One major thrust of the program is the performance measurement of local area network (LAN) hardware, performance measurement of communication protocol implementations using this hardware, and subsequent analysis of the measurements using simulation and other modeling techniques. This work will lead to realistic estimates of LAN performance in DOE environments; in addition, the enhancements to network hardware and software necessary to achieve required performance characteristics can be determined. A second major research area involves distributed computing and network access in DOE environments. The design and development of scientific workstations is the primary focus, with special emphasis upon their integration with existing local computational facilities and the Energy Research supercomputer network. The activities range from application software design for existing workstation computers to the design and implementation of a distributed file system for heterogeneous environments to software to gateway between different transport level protocols. The ultimate goal of this work is to permit scientists to use all of the computational tools at their disposal from within an environment tailored to their needs.	10/01/85	09/30/86	300	315

Summary of Applied Mathematical Sciences Research Projects
Fiscal Years 1984 and 1985

Institution	Project Manager/ Investigator(s)	Title of Project	Summary of Project in FY 1985	Start Date	End Date	FY 84 ($000)	FY 85 ($000)
Lawrence Livermore National Laboratory	Robert Huddleston/ James McGraw	Advanced Computing Concepts	Research in advanced computing concepts focuses on algorithm and language development and implementation as well as efficient programming techniques in a multitasking, multiprocessing computational environment. New research is planned in parallel algorithm design for solving multiple scales problems, in developing and implementing multitasking techniques for large scale numerical models applicable to various multitasking configurations, and in automated programming for numerical simulation systems. Research will continue in modeling and analysis of multiple scales problems, the numerical solution of PDEs arising from multiple scales problems, the numerical solution of systems of linear equations, algorithms and language development suitable for parallel and vector processors, and algorithm development for solving hydrodynamics problems.	10/01/85	09/30/86	450	555
Los Alamos National Laboratory	Norman Morse/	Advanced Computational Concepts	Serious efforts to overcome the speed limitations inherent in single processor computers have invariably involved parallel processing. The next generation of commerical supercomputers will all be shared memory parallel machines having from 8 to 16 very fast (probably vector) processors. Efficient utilization of these machines will require some significant changes in the techniques used in code development. Different algorithms will have to be created and in some cases, new mathematical models, languages, and techniques developed. In the more distant future, massively parallel machines have been proposed, consisting of hundreds and even thousands of processors to be applied to a single problem. Los Alamos has been actively involved in parallel processing for over four years, having obtained seminal results on parallel processing of applications codes. This project consists of research activities in the areas of experimental equipment, algorithm development, language research, and computational modeling methodology for parallel applications.	10/01/85	09/30/86	700	735

Summary of Applied Mathematical Sciences Research Projects
Fiscal Years 1984 and 1985

Institution	Project Manager/ Investigator(s)	Title of Project	Summary of Project in FY 1985	Start Date	End Date	FY 84 ($000)	FY 85 ($000)
Massachusetts Institute of Technology	Edward Fredkin/	Information Preserving Models of Physics and Computation	The Information Mechanics Group in MIT's Laboratory for Computer Science is developing the field of information mechanics, which deals with dynamical systems that can be thought of either as models of idealized physical processes or as physical models of computational processes. These systems are derived from conventional models of computation by adding contraints to bring the models into better agreement with physical principles; or they can be derived from conventional models of physical processes by introducing certain assumptions that are required for finitary computational models. The basic assumptions of information mechanics are: (a) the structure and the laws of dynamical systems are such that their evolution can be computed exactly by finite means, and (b) the laws that govern the evolution of these dynamical systems are exactly reversible (i. e., characterized by an invertible law, not necessarily a law that is invariant under time reversal). The second point implies that information is conserved in the evolution of the system. The investigators propose to apply these ideas to simulate physical processes involving exponential decay, turbulence, and in general, Hamiltonian dynamical systems. Simulators can be implemented with specially designed VLSI chips that give orders of magnitude speed up over general purpose computers. These ideas may provide valuable insight into the nature of reversible, discrete models of physics and computation.	07/15/83	07/14/86	110	121

Summary of Applied Mathematical Sciences Research Projects
Fiscal Years 1984 and 1985

Institution	Project Manager/ Investigator(s)	Title of Project	Summary of Project in FY 1985	Start Date	End Date	FY 84 ($000)	FY 85 ($000)
Massachusetts Institute of Technology	Jack Dennis/	High Performance Data Flow Computers	This research effort investigates the data flow approach to high performance computer systems through six closely related projects: language development, prototype evaluation facility, program translation, data structures, architecture description and simulation, and specification of a full scale data flow processor. The language development project seeks to develop a user oriented higher level language that provides a mechanism for specifying an algorithm that can be translated into a form of directed graphs. The VAL language has been developed for this purpose. The prototype evaluation facility allows the design of a machine instruction set that is suitable for representing the directed graphs of the data flow machine. The program translation project will provide tools for translating from the user language, VAL, to the machine instruction set. The data structures project will provide data structures supporting the data flow concepts and which are useful for numerical computations. The architecture description and simulation project provides tools for the analysis of possible implementations of data flow machines. The specification project will summarize conclusions about the applicability of the data flow concepts to typical large scientific and engineering codes.	01/01/85	12/31/85	231	280

Summary of Applied Mathematical Sciences Research Projects
Fiscal Years 1984 and 1985

Institution	Project Manager/ Investigator(s)	Title of Project	Summary of Project in FY 1985	Start Date	End Date	FY 84 ($000)	FY 85 ($000)
New York University, Courant Institute	Peter D. Lax/ Malvin Kalos	Advanced Computing Concepts: the Ultracomputer Project	Jack Schwartz and Alan Gottlieb are investigators on this project. This project entails the study of architectures for highly parallel computing systems, perhaps comprising many thousands of tightly coupled high performance microcomputers. Machines of this type could develop compute rates several hundreds of times larger than today's most powerful machines. Using such new machines efficiently will require developing new types of operating systems and adapting existing programming methods and languages. Effective, highly parallel coordination techniques are particularly important if efficiency is not to be lost. Machines of this power can play a central role in meeting DOE requirements for large scale computing. Significant joint research with IBM has begun on a parallel processor architecture called the RP3 that is based on ultracomputer principles. It is likely that a 16 processor prototype will appear at the beginning of 1986 and a 512 processor prototype could appear in 1987. We expect to enhance our program to respond to the opportunities implicit in this development. Specifically, we will propose to acquire one of each prototype and prepare for their use here and, through networking, by the entire DOE computational research community.	10/01/85	09/30/86	660	1188

Summary of Applied Mathematical Sciences Research Projects
Fiscal Years 1984 and 1985

Institution	Project Manager/ Investigator(s)	Title of Project	Summary of Project in FY 1985	Start Date	End Date	FY 84 ($000)	FY 85 ($000)
University of Colorado	Leon Osterweil/	Integrated Testing and Analysis of Mathematical Software Tools	The associated investigators on this project are Dieter Boecher and John Shultis. The Toolpack project that preceded this project was a joint effort of the University of Colorado, Argonne National Laboratory, University of Arizona, NAG, LTD., and Purdue University. The prototype Toolpack system developed under the original contract is being distributed by the National Energy Software Center at Argonne and by NAG, LTD. This project will build a well integrated ensemble of testing and analysis software tools based on an augmented Toolpack/IST system. Graphics tools for aiding in the analysis and debugging of mathematical software will be developed and tested. The design of graphical analysis and display tools is of particular importance to DOE for mathematical modeling software development. The research is focused on analytical displays of program development environments on scientific workstations. The complementary use of instrumentation and analysis could provide an effective framework for organizing and conceptualizing the orderly and systematic process of program verification.	08/01/84	07/31/87	175	145

Summary of Applied Mathematical Sciences Research Projects
Fiscal Years 1984 and 1985

Institution	Project Manager/ Investigator(s)	Title of Project	Summary of Project in FY 1985	Start Date	End Date	FY 84 ($000)	FY 85 ($000)
University of Illinois	David Kuck/ Duncan Lawrie	Center for Supercomputing Research and Development (CEDAR)	Ahmed Sameh and Edward Davidson are co-investigators on this project. This grant supports research in computer science and architecture, software, and algorithms and applications for the CEDAR architecture. The goals of this three year effort are to demonstrate the operation of an experimental research computer with the CEDAR architecture, complete with the optimizing, structure analyzing compiler based on Parafrase and KAP, new problem decomposition techniques, and algorithms applicable to large scale scientific computation relevant to DOE mission programs. The three components of the project are as follows. Hardware: In year one, two clusters will be purchased from a suitable vendor and connected in some suitable manner such that a two cluster system will be available for use in the project. During year two, two prototype boards will be designed and fabricated that will allow for the connection of up to eight clusters in a general way so as to support CEDAR implementation of software. By the beginning of year three the original two cluster system will be connected to provide a four cluster operational CEDAR experimental machine. During year three a larger multilevel switch capable of supporting up to 1024 processors will be designed. At this point the advisability of purchasing more clusters can be determined. Software: In year one a working prototype of the optimizing compiler based on Parafrase will be available. A production compiler based on KAP but specialized to the CEDAR architecture should be ready by the first quarter of year two. Development of the KAP based compiler will continue throughout the life of the project. Algorithms and Applications: Performance of nine applications selected by project researchers will be measured on the various versions of the CEDAR machine as the project proceeds. This will be an ongoing effort and will include work by collaborators (such as staff from Argonne) as well as the CEDAR staff.	02/15/85	02/14/88	226	2300

Summary of Applied Mathematical Sciences Research Projects
Fiscal Years 1984 and 1985

Institution	Project Manager/ Investigator(s)	Title of Project	Summary of Project in FY 1985	Start Date	End Date	FY 84 ($000)	FY 85 ($000)
University of Texas	James Browne/ Clement Leung	A Unified Approach to Parallel Computation	Components of this research project are: the construction and evaluation of a simulation modeling system for parallel computation structures; the development of an architecturally independent programming system; and investigation of a formulation support environment for parallel computation. Proposed work on systematic development of algorithmically specialized architectures may be investigated if the time and budget constraints allow. Preliminary work on the Computational Structures Language demonstrated the usefulness of abstracting parallel programming constructs in a formalism that allows analysis of algorithms in terms of computational units, communication structures, and synchronization primitives. This is a viable approach to determining the relative efficiencies of mapping equivalent algorithms devised to solve a computational problem onto various parallel architectures - that is, it provides a formal mechanism for "benchmarking" problems on different architectures without having to specify a particular language or implementation strategy a priori.	09/01/85	08/31/88	123	196

Summary of Applied Mathematical Sciences Research Projects
Fiscal Years 1984 and 1985

Institution	Project Manager/ Investigator(s)	Title of Project	Summary of Project in FY 1985	Start Date	End Date	FY 84 ($000)	FY 85 ($000)
** Subtotal **						4960	9189
** Budget and Reporting Activity: KC 070200							
Ames Laboratory, Iowa State University	James Corones/	Network Access and Research Computing Facility	Network access includes the operation and maintenance of a Vax 11/750, graphics workstation, and a link to the laboratory's Vax/FPS 164 facilities.	10/01/85	09/30/86	0	25
Argonne National Laboratory	Paul Messina/	Energy Sciences Advanced Computation	The network access activity encompasses three projects aimed at providing access to supercomputers and computer network facilities for Energy Research programs: (1) the Energy Research supercomputer network gateway facility, (2) the massively parallel supercomputer center, and (3) Argonne's research computing facility. Work has just begun on the first two activities; the research computing facility continues to provide a computing environment tailored to mathematical and computer science research.	10/01/85	09/30/86	100	225
Brookhaven National Laboratory	Ronald Peierls/	Network Access	Network access includes a 50 KB link to the NYU TAC.	10/01/85	09/30/86	0	25

Summary of Applied Mathematical Sciences Research Projects
Fiscal Years 1984 and 1985

Institution	Project Manager/ Investigator(s)	Title of Project	Summary of Project in FY 1985	Start Date	End Date	FY 84 ($000)	FY 85 ($000)
Florida State University	Joseph Lannutti	Supercomputer Computations Research Institute	The Supercomputer Computations Research Institute conducts research in computational sciences, including physics, chemistry, geophysics, oceanography, applied mathematics, and others. Supercomputer access is provided to on site researchers and to other qualified OER researchers through the national NMFECC network.	10/01/84	09/30/89	0	7000
Lawrence Berkeley Laboratory	Carl Quong/ Joseph Sventek	Network Access	Network access activity provides the computational resources needed to support the Laboratory's basic research program in Applied Mathematical Sciences. Maintenance and improvement of the local research computing facilities which provide the standard functions such as program development, electronic mail, network access, document preparation, text processing, high quality graphics, etc., will continue. Additionally, special attention will be directed toward the problems of remote access to the ER supercomputer network through local computational facilities which provide an integrated computing environment to the user.	10/01/85	09/30/86	100	100
Lawrence Livermore National Laboratory	John Killeen/	National Magnetic Fusion Energy Computer Center and Network	This reflects the portion of the NMFECC budget that provides supercomputer access to the non-MFE researchers in ER, including network access, upgrade of ports at ER sites (SLAC, Fermi, ANL, LBL, and others).	10/01/85	09/30/86	100	5943

Summary of Applied Mathematical Sciences Research Projects
Fiscal Years 1984 and 1985

Institution	Project Manager/ Investigator(s)	Title of Project	Summary of Project in FY 1985	Start Date	End Date	FY 84 ($000)	FY 85 ($000)
Los Alamos National Laboratory	Billy Buzbee/ David Campbell	Workshops in Computing and Mathematical Sciences	Each year the Los Alamos Computing Division and the Center for Nonlinear Studies host a series of workshops addressing leading issues in computing and mathematical sciences. The purpose of this program is to help defray associated costs. Costs of these workshops include publication (agendas, abstracts, proceedings), and travel support for speakers.	10/01/85	09/30/86	100	100
New York University, Courant Institute	Peter D. Lax/ Max Goldstein	Research Computing Facility and Network Access	The research computing facility provides the computational resources needed to support the ongoing basic research in Applied Mathematical Sciences. It serves as a testbed for the evaluation and deployment of new computing capabilities. Maintenance and development of this facility, including integration and evaluation of emerging computing technologies, is a focal point for collaborative interaction with other DOE research sites and with researchers in computer science throughout the university and industrial communities. The experimental effort places special emphasis on issues related to distributed computing environments and the applicability of varied computational resources to DOE scientific researchers. The goals of the experimentation are to: (a) provide a variety of computer facilities to the CMCL community; (b) acquire and devise means to easily enable communications among the various hardware and software elements comprising the research facility; and (c) discover and implement means that reduce the impact on the user of the complexity inherent in multicomputer environments.	10/01/85	09/30/86	150	150

Summary of Applied Mathematical Sciences Research Projects
Fiscal Years 1984 and 1985

Institution	Project Manager/ Investigator(s)	Title of Project	Summary of Project in FY 1985	Start Date	End Date	FY 84 ($000)	FY 85 ($000)
Oak Ridge National Laboratory	Robert Ward/	Research Computing Facility and Network Access	The network access task provides for the improvement, operation, and maintenance of a research computing facility for the Mathematical Sciences section. The principal item in our facility is a DEC Vax 11/780 acquired late in FY 1983. Connection to the Defense Data Network (Milnet/Arpanet) took place in 1984. This facility will continue to be upgraded and be used to explore innovative approaches to scientific computing and to communicate with researchers at other DOE facilities. It is anticipated that this facility will also serve as the network gateway to an experimental prototype parallel computer to be acquired in FY 1985.	10/01/85	09/30/86	75	100
Pacific Northwest Laboratory/ Battelle	Marvin Erickson/ Wesley Nicholson	Network Access and Research Computing Facility	The research computing facility is operated and maintained for use by Applied Mathematical Sciences researchers.	10/01/85	09/30/86	0	100

STRATEGIC COMPUTING

New-Generation Computing Technology:
A Strategic Plan for its Development
and Application to Critical
Problems in Defense

Defense Advanced Research
Projects Agency

28 October 1983

EXECUTIVE SUMMARY

To meet the challenge of certain critical problems in defense, the Defense Advanced Research Projects Agency (DARPA) is initiating an important new program in Strategic Computing. By seizing an opportunity to leverage recent advances in artificial intelligence, computer science, and microelectronics, the Agency plans to create a new generation of "machine intelligence technology." This new technology will have unprecedented capabilities and promises to greatly increase our national security and our economic strength as it emerges during the coming decade.

THE CHALLENGE. Computers are increasingly employed in defense, and are relied on to help us hold the field against larger forces. But current computers, having inflexible program logic, are limited in their ability to adapt to unanticipated enemy behavior in the field. We are now challenged to produce adaptive, intelligent systems having capabilities far greater than current computers, for use in diverse applications including autonomous systems, personalized associates, and battle management systems. The new requirements severely challenge the technology and the technical community.

THE OPPORTUNITY. Within the past few years, important advances have occurred in many separated areas of artificial intelligence, computer science, and microelectronics. Advances in "expert system" technology now enable the mechanization of the practical knowledge and the reasoning methods of human experts in many fields. Advances in machine vision, speech, and machine understanding of natural language provide easy ways for humans to interact with computers. New ways to structure the architectures of computers enable computations to be processed in parallel, leading to large improvements in machine performance. Finally, new methods of microsystem design and implementation enable the rapid transfer of new architectural concepts into state-of-the-art microelectronics.

These separate advances can be jointly exploited to mechanize the thinking and reasoning processes of human experts into the form of powerful computing structures implemented in microelectronics, thus creating machine intelligence technology of unprecedented capabilities. The new requirements for adaptive intelligent military systems serve to integrate activities in the separate areas shown in Table 1 and guarantee the leveraging of the key advances.

TABLE 1. KEY AREAS OF ADVANCES THAT CAN BE LEVERAGED TO PRODUCE HIGH-PERFORMANCE MACHINE INTELLIGENCE

- Expert Systems: Codifying and mechanizing practical knowledge, common sense, and expert knowledge
- Advances in Artificial Intelligence: Mechanization of speech recognition, vision, and natural language understanding.
- System Development Environments: Methods for simplifying and speeding system prototyping and experimental refinement
- New Theoretical Insights in Computer Science
- Computer Architecture: Methods for exploiting concurrency in parallel systems
- Microsystem Design Methods and Tools
- Microelectronic Fabrication Technology

GOALS AND METHODS. The overall goal of the Strategic Computing Program is to provide the United States with a broad line of machine intelligence technology and to demonstrate applications of the technology to critical problems in defense. Figure 1 provides a summary overview of the program structure and goals.

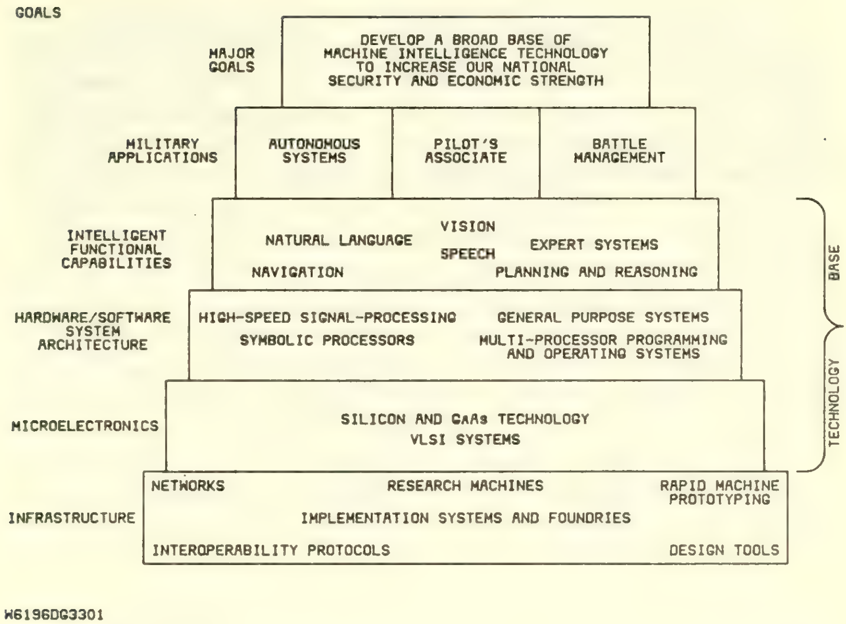


FIGURE 1. PROGRAM STRUCTURE AND GOALS

The program begins by focusing on demanding military applications that require machine intelligence technology. The applications generate requirements for functions such as vision, speech, natural language, and expert system technology, and provide an experimental environment for synergistic interactions among developers of the new technology. The intelligent functions will be implemented in advanced architectures and fabricated in microelectronics to meet application performance requirements. Thus, the applications serve to focus and stimulate or "pull" the creation of the technology base. The applications also provide a ready environment for the demonstration of prototype systems as the technology compartments successfully evolve. To carry out this program, DARPA will fund and coordinate research in industrial, university, and government facilities, and will work with the Military Services and Defense Agencies to insure successful transfer of the resulting technology.

Figure 2 provides an overview of program activity and suggests ways of visualizing and interpreting how the various compartments of program activity will unfold over time.

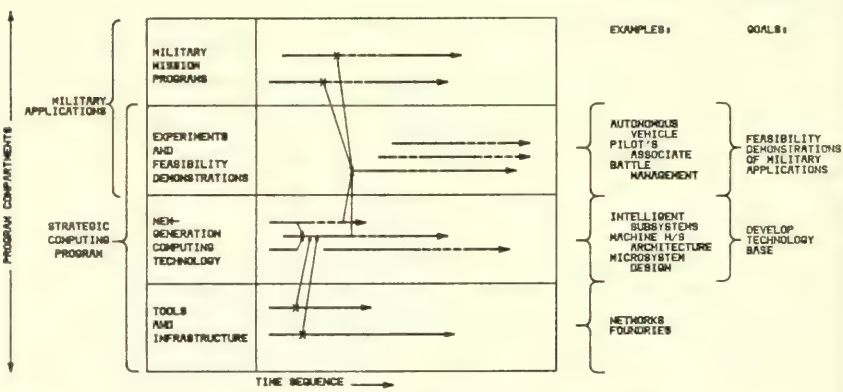


FIGURE 2. VISUALIZING & INTERPRETING THE PROGRAM'S COMPARTMENTS, ELEMENTS, AND TIMELINE

ACTIVITIES AND PLANS. The initial program applications include an autonomous vehicle, a pilot's associate and a carrier battle group battle management system. These applications stress different compartments of machine intelligence technology, and exert a strong pull on the overall technology. These specific examples were selected for inclusion in the Strategic Computing program based on a weighted consideration of the following factors:

- The application must effectively employ the new technology to provide a major increase in defense capability in light of realistic scenarios of combat situations that might occur at the future time when the new systems can be procured and deployed.
- The application must provide an effective "pull" on the new generation technology. It must demand an aggressive but feasible level of functional capability from one or more of the intelligent functions at appropriate points in the timeline.
- Development of the application must lead to new engineering know-how in artificial intelligence software areas, such as planning and reasoning, learning, navigation, knowledge base management, and so on.
- The application must test the efficacy of the new technology at a realistic quantitative scale of performance demands. In this way we seek to ensure against unexpected quantitative changes in system performance as a result of scaling up from models and laboratory experiments to real systems.
- The application must provide an effective experimental "test-bed" for evolving and demonstrating the function(s). Stability over time, access, and visibility are thus important factors.
- The application must effectively leverage program resources. Thus an important factor is the extent to which an existing military program provides a base of capital resources and experienced

personnel into which the new generation technology can be experimentally introduced (versus this program having to provide such non-computing resources).

- It is important to choose a mix of applications that are jointly supportive of and involve all three Services, and which are appropriately executed through each Service. Only in this way can we develop the base for extension of this technology into a wide range of military systems.
- Finally, an important selection factor is the potential provided by the specific application for effecting the transfer into the services of the new machine intelligence technology.

The planning timelines for evolving these applications have been interlocked with program timelines for evolving intelligent functions (such as machine vision, speech, and expert system technology). The plans for creating machine intelligence capabilities have in turn been interlocked with the program plans for system architectures that support the signal processing, symbolic processing, and general-purpose processing underlying the machine intelligence.

The planned activities will lead to a series of demonstrations of increasingly sophisticated machine intelligence technology in the selected applications as the program progresses. Milestones have been established for the parallel development of the machine architectures required to support these demonstrations.

Attention is focused early in the program on provision of the necessary infrastructure to support and coordinate the activities of the many people and organizations that will be involved in the program. Computing facilities, network services, interoperability standards, access to rapid system prototyping and integrated circuit implementation services must all be in place for the enterprise to succeed. This will also insure rapid propagation of the knowledge and technology produced by the program into the community of participants and into US industry.

MANAGEMENT AND FUNDING. Management of the Strategic Computing Program will be carried out by the Defense Advanced Research Projects Agency. Within DoD, DARPA will coordinate closely with USDRE and the Military Services. A Defense Science Board panel has been convened to make recommendations on DoD utilization of machine intelligence technology. Other advisory panels and working groups will be constituted, with representatives from industry, universities, and government, to provide additional required advice in specific areas.

Table 2 shows the annual cost for the Strategic Computing Program. Program costs for the first five years of the program are estimated to be approximately 600 million dollars. The logic of the sequencing of activities is reflected in the breakdown of spending in the first three categories. Relative spending on tools and infrastructure is higher early in the program. The large technology base activity and component of spending will likely peak in FY 87-88. Applications activity and spending expand moderately at first, then rapidly in the late 80s, peaking near the end of the program. The entire program will peak about the end of the decade, declining thereafter as program goals are achieved.

TABLE 2. STRATEGIC COMPUTING COST SUMMARY IN $M
(* Out-year funding levels to be determined by program progress.)

	FY84	FY85	FY86	FY87*	FY88*
Total Military Applications	6	15	27	TBD	TBD
Total Technology Base	26	50	83	TBD	TBD
Total Infrastructure	16	27	36	TBD	TBD
Total Program Support	2	3	4	TBD	TBD
TOTAL	50	95	150	TBD	TBD

53-831 O -86 - 16

The basic acquisition policy is that military applications will be carried out by industry drawing upon results of research carried out in the universities. Advanced computer architectures will be developed primarily in joint projects between universities and industry. Most of the hardware and software development efforts will be competed. The most advanced artificial intelligence ideas that seem ripe for development will be exploited with heavy university involvement. For these, expert judgment from leading participants in the field will be sought and directed selection will result. Construction and access to computing technology infrastructure will be competed.

We intend a significant effort toward technology transfer of results of this program into the military services. This effort will include: (a) use of Service Agents and Service COTRs; (b) a process of cost-sharing with the Services in the development of military applications; (c) the inclusion of technology base results from this program in Service Programs and Testbeds, and (d) training of Service personnel by involvement in technology base developments.

Equally important is technology transfer to industry, both to build up a base of engineers and system builders familiar with computer science and machine intelligence technology now resident in leading university laboratories, and to facilitate incorporation of the new technology into corporate product lines. To this end we will make full use of regulations for Government procurement involving protection of proprietary information and trade secrets, patent rights, and licensing and royalty arrangements.

Communication is critical in the management of the program, since many of the important contributors will be widely dispersed throughout the US. Unique methods will be employed to establish a productive research community and enable participants to interact with each other and to interlock with the program plan. Existing computer tools such as electronic networks and message systems will be used to coordinate program activities. More advanced methods will include provision to participants of remote electronic views of, and interactions with, the evolving program planning timelines.

CONCLUSIONS. We now have a plan for action as we cross the threshold into a new generation of computing. It is a plan for creating a large array of machine intelligence technology that can be scaled and mixed in countless ways for diverse applications.

We have a plan for "pulling" the technology-generation process by creating carefully selected technology interactions with challenging military applications. These applications also provide the experimental test beds for refining the new technology and for demonstrating the feasibility of particular intelligent computing capabilities.

The timely, successful generation and application of intelligent computing technology will have profound effects. If the technology is widely dispersed in applications throughout our society, Americans will have a significantly improved capability to handle complex tasks and to codify, mechanize, and propagate their knowledge. The new technology will improve the capability of our industrial, military and political leaders to tap the nation's pool of knowledge and effectively manage large enterprises, even in times of great stress and change.

Successful achievement of the objectives of the Strategic Computing initiative will lead to deployment of a new generation of military systems containing machine intelligence technology. These systems will provide the United States with important new methods of defense against massed forces in the future - methods that can raise the threshold and diminish the likelihood of major conflict.

There are difficult challenges to overcome in order to realize the goals of a national program of such scope and complexity. However, we believe that the goals are achievable under the logic and methods of this plan, and if we seize the moment and undertake this initiative, the Strategic Computing Program will yield a substantial return on invested resources in terms of increased national security and economic strength.

TABLE OF CONTENTS

CHAPTER 1
INTRODUCTION

As a result of a series of advances in artificial intelligence, computer science, and microelectronics, we stand at the threshold of a new generation of computing technology having unprecedented capabilities. The United States stands to profit greatly both in national security and economic strength by its determination and ability to exploit this new technology.

Computing technology already plays an essential role in defense technologies such as guided missiles and munitions, avionics, and C^3I. If the new generation technology evolves as we now expect, there will be unique new opportunities for military applications of computing. For example, instead of fielding simple guided missiles or remotely piloted vehicles, we might launch completely autonomous land, sea, and air vehicles capable of complex, far-ranging reconnaissance and attack missions. The possibilities are quite startling, and suggest that new generation computing could fundamentally change the nature of future conflicts.

In contrast with previous computers, the new generation will exhibit human-like, "intelligent" capabilities for planning and reasoning. The computers will also have capabilities that enable direct, natural interactions with their users and their environments as, for example, through vision and speech.

Using this new technology, machines will perform complex tasks with little human intervention, or even with complete autonomy. Our citizens will have machines that are "capable associates," which can greatly augment each person's ability to perform tasks that require specialized expertise. Our leaders will employ intelligent computers as active assistants in the management of complex enterprises. As a result the attention of human beings will increasingly be available to define objectives and to render judgments on the compelling aspects of the moment.

A very broad base of existing technology and recent scientific advances must be jointly leveraged in a planned and sequenced manner to create this new intelligent computer technology. Scientists from many disciplines, scattered throughout the universities, industry, and government must collaborate in new ways, using new tools and infrastructure, in an enterprise of great scope. Adaptive methods of planning must be applied to enhance the process of discovery. Events must be skillfully orchestrated if we are to seize this opportunity and move toward timely success.

In response to these challenges and opportunities, the Defense Advanced Research Projects Agency (DARPA) proposes to initiate an important new program in Strategic Computing. To carry out this program, DARPA will fund and coordinate research in industrial, university, and government facilities, and will work with the Military Services and Defense Agencies to insure successful transfer of the results.

The overall goal of the program is to create a new generation of machine intelligence technology having unprecedented capabilities and to demonstrate applications of this technology to solving critical problems in Defense. Although the achievements of the program applications' objectives will significantly improve the nation's military capabilities, the impact of nonmilitary spin-offs on the national economy should not be underestimated. This document provides an overview of the proposed program.

CHAPTER 2
THE MILITARY CHALLENGE

Adaptive Technology Is Important to Defense. Computers are being increasingly employed to support United States military forces. The growing complexity of forces and rising level of threats have stimulated the use of ever more advanced computers. Improvements in the speed and range of weapons have increased the rate at which battles unfold, resulting in a proliferation of computers to aid in information flow and decision making at all levels of military organization. Smarter computerized weapons and forces are now depended upon to be able to hold the field against superior numbers.

A countervailing effect on this trend is the rapidly decreasing predictability of military situations, which makes computers with inflexible logic of limited value. Consider a problem encountered with a current generation computerized system during a recent conflict. A radar designed to automatically acquire and track aircraft was supposed to follow all aircraft maneuvers, recognize countermeasures, and not become confused or lose track because of dropped decoys. The other side (who also had the same radar) innovated the tactic of approaching in groups of four aircraft and, as they came over the horizon, rapidly branching (fleur dé lis) into four different directions. The computer controlled radar reacted by jittering around the centroid until it lost all four tracks.

The solution to this unforeseen problem is simple from a logical viewpoint, but there was no way for the forces in the field to codify and implement the solution. Instead, once the problem was recognized and diagnosed in the field, an equivalent situation was created and the solution was programmed and evaluated in the homeland, and the new software/firmware was then flown to the radar locations in the field. Even with a crash program, it took several days to eliminate the radar's inflexibility when responding to a simple change of tactics that had not been anticipated by the radar designers.

Confronted with such situations, leaders and planners will continue to use computers for routine tasks, but will often be forced to rely solely on their people to respond in unpredictable situations. Revolutionary improvements in computing technology are required to provide more capable machine assistance in such unanticipated combat situations. The military requirements for dealing with uncertainty and information saturation in life-threatening situations are far more demanding of the technology than evolving needs in the civilian sector.

Intelligent Military Systems Demand New Computer Technology. The effects of increasing unpredictability are evident over a wide range of military computer applications. In certain routine military tasks -- surveillance, monitoring, and recording systems -- computers have actually replaced human operators. Small scale computer systems have been applied in precision guided munitions ("smart weapons") and some reconnaissance devices. To achieve truly autonomous systems, a variety of complex functions must be performed. However, the emergence of autonomous systems is inhibited by the inability of present computers to robustly direct actions that fulfill mission objectives in unpredictable situations. Commanders remain particularly concerned about the role autonomous systems would play during the transition from peace to hostilities when rules of engagement may be altered quickly.

An extremely stressing example of such a case is the projected defense against strategic nuclear missiles, where systems must react so rapidly that it is likely that almost complete reliance will have to be placed on automated systems. At the same time, the complexity and unpredictability of factors affecting decisions will be very great.

In many military activities, people are often saturated with information requiring complex decisions to be made in very short times. This is a severe problem for operators of complex combat systems such as aircraft, tanks, and ships. The physical environment -- noise, vibration, and violent maneuvers -- is extremely taxing; moreover, the information flowing to the operator increases dramatically as missions become more demanding, sensor and weapons systems become more complex, and threats to survival

become more numerous and serious. The ability of computers to assist in such situations is limited because their computational capability cannot handle these highly complex unstructured environments; in addition, their interface with humans is so (cognitively) inefficient that it is doubtful the person could receive, interpret, and act on the information in time even if it were available within the machine. Improvements can result only if future computers can provide a new "quantum" level of functional capabilities.

The management of large-scale military enterprises requires large staffs to gather information, to develop and evaluate alternative courses of action, and to construct detailed plans. The trend in all areas toward faster-moving warfare severely stresses the whole staff function. Greater uncertainty in the military environment forces consideration of more options. Increasingly sophisticated methods of deception, countermeasures, and camouflage make timely acquisition of vital information more difficult. Improved weapon speed and range increase the scale of military actions. The result is a growing uncertainty in the decision making process and the evolution of large, labor-intensive military command organizations. Current computers provide only limited assistance to such decision making because they have limited ability to respond to unpredictable situations and to interact intelligently with large human staffs.

Across this spectrum of applications, from autonomous systems to systems aiding in battle management, we need computers that have far more capability for intelligent operation, improved survivability in hostile and high-radiation environments, and greatly improved man-machine interfaces. Many isolated pieces of the required technology are already being developed. The challenge is to exploit these beginnings, make new efforts to develop the full set of required technologies, and integrate components of the emerging new technology in order to create revolutionary defense capabilities. Such revolutionary capabilities can provide our nation with new, highly flexible and significantly improved defenses against possible assaults by massed forces in the future.

CHAPTER 3
THE TECHNICAL OPPORTUNITY

A New Generation of Computing Technology. Within the past few years, a series of important advances have occurred across a wide range of areas in artificial intelligence, computer science, and microelectronics. By jointly leveraging these many separate advances, it will be possible to create a completely new generation of machine intelligence technology having unprecedented capabilities.

KEY AREAS OF ADVANCES THAT CAN BE LEVERAGED TO PRODUCE HIGH-PERFORMANCE MACHINE INTELLIGENCE

- o Expert Systems: Codifying and mechanizing practical knowledge, common sense, and expert knowledge
- o Advances in Artificial Intelligence: Mechanization of speech recognition, vision, and natural language understanding.
- o System Development Environments: Methods for simplifying and speeding system prototyping and experimental refinement
- o New Theoretical Insights in Computer Science
- o Computer Architecture: Methods for exploiting concurrency in parallel systems
- o Microsystem Design Methods and Tools
- o Microelectronic Fabrication Technology

Advances in microelectronic technology have led to the manufacturability of silicon integrated-circuit chips consisting of hundreds of thousands of transistors. Because of their tiny size, the transistors in such chips function at very high switching speeds and consume very little power. New methods of microsystem design enable designers to rapidly design and implement digital systems in microelectronics.

Computer scientists have developed new insights into how the exploitation of area, time, and energy tradeoffs in computing systems. This work assures the feasibility of radically new forms of computing structures.

These advances in theory are now guiding computer architects in their search for ways to exploit concurrency in highly parallel systems. A number of research groups have produced workable concepts for such machines; these concepts include methods for achieving parallelism in machines that provide very high performance processing on unstructured, complex problems. Advances in system development environments now enable very rapid prototyping of hardware and software for such new machines.

Perhaps the most stunning advances have come in the area of expert systems. The term "expert system" describes the codification of any process that people use to reason, plan, or make decisions as a set of computer rules. For example, a detailed description of the precise thought processes and heuristics by which a person finds their way through a city using a map and visual landmarks might be codified as the basis of an "expert system" for local navigation. The methods for identifying and mechanizing practical knowledge, common sense, and expert knowledge have solidified and are now finding wide application. Expert systems, mechanized at the level of practical reasoning, now stand in great contrast to systems created using traditional computing technology. Rather than being "black boxes" whose internal workings are inaccessible to users, these systems have the ability to "explain" the reasoning used to reach decisions or take actions. The knowledge base that guides their operation can be changed quickly to cope with changes in the environment, thereby easing adaptation to situations like the "radar problem" described earlier. The methods of programming such systems promise to stimulate a movement towards articulating, codifying, and better exploiting a wide range of practical human knowledge.

Finally, there have been very important successes in other areas of artificial intelligence (AI), particularly in the mechanization of vision and visual-motor interaction, the mechanization of speech recognition, and in the understanding of natural language (see Chapter 5, Section 5.2.1).

Form and Functions of Machine Intelligence Technology. Properly combined, all these recent advances now enable us to move toward a completely new generation of machine intelligence technology. What kind of special

capabilities would the new computers have? First, they would be able to perform intelligent functions such as:

- o UNDERSTANDING NATURAL LANGUAGE EXPRESSIONS
- o INFORMATION FUSION AND MACHINE LEARNING
- o PLANNING AND REASONING

They would also be able to interact with their users and environment through natural modes of sensory communication such as:

- o VISION AND VISUAL IMAGE GENERATION
- o SPEECH RECOGNITION AND PRODUCTION

What form might these machines take? One important characteristic is that instead of being a single collection of microelectronics which fills all needs, the new generation of "intelligent" computer systems will be modular (conceptually; even if not physically in all cases). Each system will be created by combining modules from different specialized compartments of the new technology base, much as one might now compose a "component video system."

For example, consider a small modular computer system used to control a future autonomous vehicle. Vision modules will be included that provide basic scene-processing and object-recognition capabilities. With vision modules as input devices, a symbolic processor module would be then able to directly process fragments of pictorial, graphic, and three-dimensional scenic images. When further supported by rule-based inferencing and image understanding in a compact but powerful symbol processor and interfaced with specialized motor-control systems, these vision modules will enable the computer-controlled autonomous vehicle to "see," to move about, and to interact "intelligently" with its environment. The resulting vision, scene interpretation, and motor control processes will be, at the very least, analogous to those found in lower animals.

This simple sketch merely hints at the possibilities. The magnitude of the opportunity before us will be seen as we explore in this document the interaction between the new-generation technology base and demanding areas of military application of machine intelligence. The "envelope" of possibilities includes not only smart autonomous systems, but also intelligent machines employed as "personal associates" to boost the capabilities of people to perform complex tasks, and as active contributors serving leaders and teams of people in the management of large enterprises.

Spinoffs from the Technology Base Can Stimulate National Economy. In addition to the planned military applications discussed in this document, the value of future commercial products made available by development of the new generation technology will be enormous. The effects will be analogous to those resulting from the replacement of the vacuum tube by the transistor, the displacement of discrete transistors by integrated circuits, and the fourth generation displacement of simple integrated circuit technology by VLSI now occurring in the computer and electronics industry.

The Strategic Computing Program promises the production of machine intelligence technology that will enable yet another major cycle of new economic activity in the computer and electronics industry. If the United States aggressively competes to develop these systems, it will gain access to enormous new commercial markets that will build on top of the successes of fourth generation technology. Spinoffs from a successful Strategic Computing Program will surge into our industrial community. They will be used by the computer industry as it creates and exploits a host of new markets for the underlying machine intelligence hardware and software technology and by the automotive and aerospace industries as they integrate intelligent CAD into the development process and intelligent CAM and robotics into manufacturing. The consumer electronics industry will integrate new-generation computing technology and create a home market for applications of machine intelligence. In addition, a wide range of service industries will emerge that create and provide new applications for machine intelligence and new ways to leverage the production, codification, and mechanization of useful human knowledge.

CHAPTER 4
GOALS AND METHODS

In response to the challenging opportunity for creating and exploiting intelligent computing technology, the Defense Advanced Research Projects Agency proposes to initiate this important new program in Strategic Computing. The overall goal of the program is to provide the United States with a broad base of machine intelligence technology that will greatly increase our national security and economic power. This technology promises to yield strong new defense systems for use against massed forces, and thus to raise the threshold and decrease the chances of major conflict.

To achieve this goal, a wide range of present technology and recent scientific advances must be leveraged in a coordinated manner. Engineers and scientists from many disciplines must collaborate in new ways in an enterprise of very large scope. A framework must be created for the effective, adaptive planning of the discovery and development processes in this enterprise. A skillful orchestration of events and exploitation of available infrastructure will be required to insure a timely success.

This chapter sketches the methods the program will use to adaptively select and schedule program activities. The chapter discusses near-term planning tactics and the plans for leveraging the interaction between selected military applications and the evolving technology base. The chapter ends with an overview of how to visualize the program's adaptive planning process and timelines.

Figure 4.1 shows the logical structure of the Program and its goals. The overall goals will be reached by focusing on three specific military applications to develop a new technology base. In order to conduct successful military demonstrations of these applications it will be necessary to develop new machine intelligence functional capabilities. Although these intelligent capabilities are largely provided by software, they depend strongly on the underlying hardware architectures for high performance and efficiency. Finally, the program depends on the exploitation of

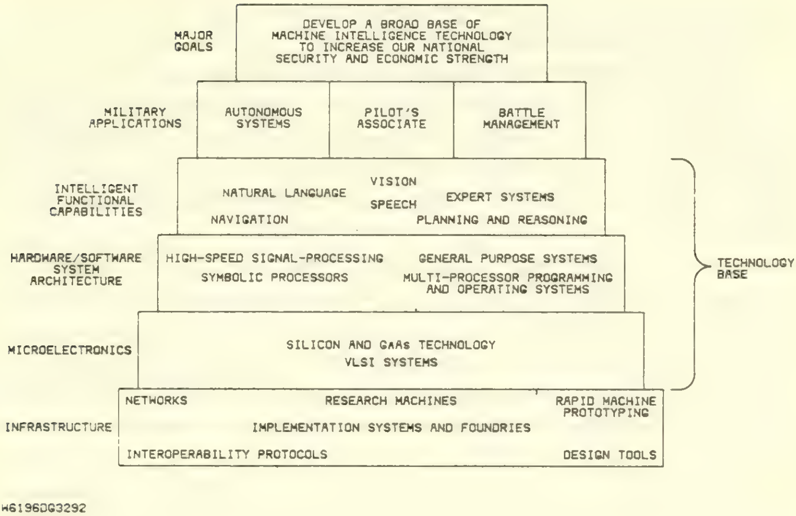


FIGURE 4.1 PROGRAM STRUCTURE AND GOALS

faster, denser, more radiation resistant, lower-power devices provided by state-of-the-art microelectronics.

Developing and Demonstrating Applications To Defense. Our projections for military applications of new generation computers cover a wide spectrum of activities. They range from applications of individual machines without operators or users as, for example, in autonomous systems, to applications involving groups of machines and groups of people collectively engaged in complex tasks, as for example in battle management.

Across this spectrum of applications we find a range of requirements for machine intelligence technology. Some autonomous systems require low-power systems, moderate performance planning and reasoning, and very powerful vision systems. At the other extreme, certain battle management systems will require immense planning and reasoning processors, vast knowledge and database management systems, perhaps no vision systems, but highly complex distributed, survivable communications systems.

Specific applications are to be identified, selected, developed, and demonstrated as discussed in Chapter 5. The applications are selected for their relevance to critical problems in Defense, and for their suitability in exerting an effective "pull" on the new generation technology base in such a way as to enable a much broader range of applications.

Creating the Technology Base. Although there is a very wide range of possible applications of the machine intelligence technology, the technology base will have many elements that are commonly used in many applications. By studying many specific applications, we have developed a taxonomy of possible future intelligent systems and identified the common functions required to create those systems. For example, many future military applications will require vision, speech, hearing, and natural language understanding functions to facilitate easy communication between people and machines. We plan to develop these common functions as modular "intelligent subsystems," and we have evolved an initial set of technical requirements for these subsystems by detailed study of specific applications described in Chapter 5 (Section 5.1) and in Appendix I.

Some of the intelligent subsystem functions, such as speech and vision, have value in a host of military and commercial systems, and generic or general purpose software and hardware can be developed independent of the application. Other intelligent functions, such as planning and reasoning (as done for example using expert systems), and information fusion (including future extensions to include systems that learn from experience) depend strongly on and must be designed for each specific application.

The development of advanced machine architectures will accompany the development of associated software to produce integrated intelligent subsystems. The development of machine architectures will be directed toward maximizing the functional power and the speed of computation. Powerful, efficient intelligent processors, database machines, simulation and control systems, display systems, and general purpose systems will be needed to achieve the program performance goals and to support selected military applications. During the early years of the program, we will investigate, refine and perfect specific computer architectures. Exploratory development, testing, and evaluation of the machines will be done in parallel with the work on software and microelectronics technology. Specific candidate architectures will then be selected for full-scale development, with their scale and configuration determined by the requirements of unfolding experimental applications.

To meet applications constraints and requirements for performance, weight, volume, power dissipation, and cost, the machine architectures will be implemented, at least initially, in advanced silicon microelectronics. The technology is widely available in industry, and accessible through implementation service infrastructure, due in part to the success of DoD VLSI/VHSIC programs. Later in the program, gallium arsenide microelectronics technology will be exploited for high performance in critical defense applications that require both low power and radiation hardness.

Thus, we aim to create

-- Integrated Intelligent Subsystems, composed of
-- Machine Hardware/Software Architectures, and
-- Microelectronics, built using
-- Tools and Infrastructure.

This last list, in fact, represents the hierarchy of development areas addressed in this plan. Specific objectives have been established for program activities in each of the technology-base areas in order to provide the functional capabilities required in the intelligent subsystems used in selected applications programs. These objectives then establish requirements for the tools and infrastructure used to support program activity.

An analysis of the technical specialties required over the long-term for this program reveals personnel shortfalls in the areas of artificial intelligence and VLSI system architecture. Efforts must be made to increase the supply of trained talent in these fields. We will encourage the offering of appropriate university courses and will encourage industry to support this process through grants, liberal re-training programs, and loan of key technical personnel for teaching. An important long-term effect of the program's technology-base development should be an increase in qualified faculty and graduate students active in all the related fields of study. This program's research is highly experimental, and significant advances require adequate computing facilities. We plan to ensure that adequate computing resources are made available to research personnel to carry out the proposed work.

Program Methodology. The program begins by building on a selected set of intelligent computing capabilities that are ripe for development in the near term. The program will develop these capabilities and accumulate further intelligent capabilities under the "pull" of demanding military applications. The objective is to evolve these capabilities into a broad base of new generation technology and to demonstrate specific applications of the new technology to solving a number of critical problems in Defense.

Artificial intelligence already offers moderately developed functional capabilities in the areas of machine vision, speech recognition, and understanding of natural language. Expert systems that perform as well as capable humans at situation analysis have already been demonstrated.

Through an analysis of numerous potential military applications of machine intelligence, an initial list of intelligent functional capabilities was developed that have common utility across many applications (see Table 4-1). Substantial progress has already been made in the development of some of these, such as speech recognition, but others such as information fusion are still in an early stage of their evolution.

Table 4-1. Improvements in Functional Capabilities to be Provided by the New-Generation of Computing Technology

Areas for major improvements in machine intelligence (processing and memory)	Areas for major improvements in interfacing machines to their users and environment (input and output)
UNDERSTANDING OF NATURAL LANGUAGE	VISION
SIGNAL INTERPRETATION	GRAPHICS DISPLAY/IMAGE GENERATION
INFORMATION FUSION/MACHINE LEARNING	SPEECH RECOGNITION AND PRODUCTION
PLANNING AND REASONING	DISTRIBUTED COMMUNICATIONS
KNOWLEDGE AND DATA MANAGEMENT	.
SIMULATION, MODELING, AND CONTROL	.
NAVIGATION	.
.	
.	
.	

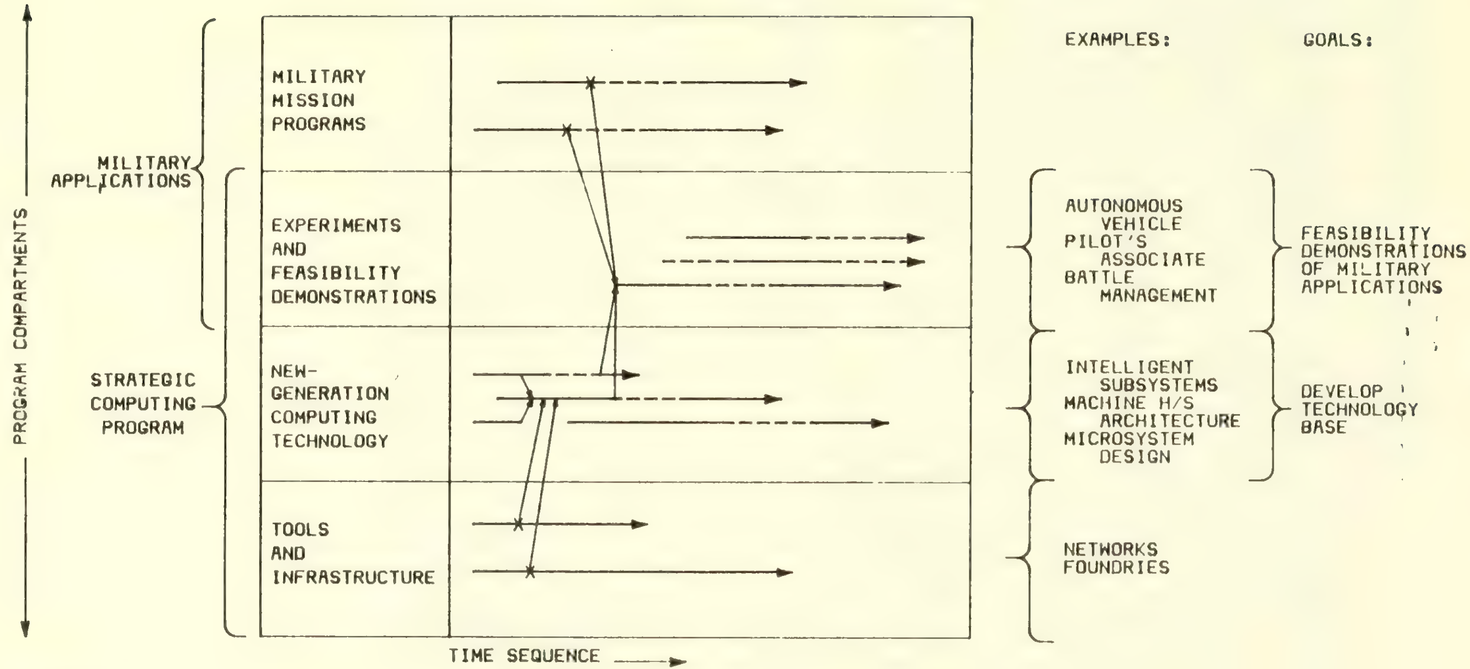
These initial functional capabilities can be scaled and combined in many ways to create a large "envelope" of intelligent systems in the future. The possibilities increase as we add new functions to the list.

A very large envelope of future military applications is also envisioned for new generation computing technology. Even if we restrict our attention to a few areas such as autonomous systems, personal associates, and computational aids for managing large enterprises, the set of possibilities is large.

It is important to note that any improvements in machine intelligence technology capabilities expand the envelope of possible applications. But how do we focus on specific capabilities to "push" at particular times? How do we select specific applications to "pull" the technology? The key is an integrated planning framework - an active planning timeline - that derives realistic, near-term application goals from credible technology developments and simultaneously stresses that technology development by proper selection of application demonstrations to focus the R&D. That process has been used in developing this plan, based on our best understandings at this time, and is described in the following sections. As technology is developed, the situation and thus the plan will change, so this should be viewed as a dynamic process.

<u>Visualizing Program Compartments and Planning Timeline</u>. Figure 4-2 provides an overview of program activity and suggests ways of visualizing and interpreting how the various compartments of program activity will unfold over time. The figure is intended to help readers interpret more detailed plans and charts, and figures that follow later in this document.

The program goals can be visualized in the figure as guiding the establishment of specific objectives for applications, experiments, and demonstrations, and specific objectives for new generation technology. Note that the Strategic Computing program intersects with military mission programs in the area of applications experimentation and demonstrations. Activities in this area of overlap are based on opportunities in military mission areas <u>and</u> opportunities in the new generation technology base, as earlier objectives are achieved in each of these areas. The figure also illustrates the role of support tools and infrastructure, and suggests how the achievement of technology base objectives depends on achievements in tool and infrastructure construction.



H6196AP3200

FIGURE 4.2 VISUALIZING & INTERPRETING THE PROGRAM'S COMPARMENTS, ELEMENTS, AND TIMELINE

CHAPTER 5
ACTIVITIES AND PLANS

We now describe the activities and plans that will be used to achieve the goals and objectives of the Strategic Computing Program. For purposes of exposition we present the key ideas by showing how example activities proceed under the plan. These examples are selected to convey the very large scope of the enterprise, and to illustrate the methods of orchestration that will be used to stimulate activity, provide joint leveraging of technologies, and direct program elements toward overall program goals.

The chapter is divided into four sections. Section 5.1 gives examples of planned application experiments and demonstrations. These applications drive new generation technology requirements that are reflected in Section 5.2, which describes how a mix of technology requirements will be provided under planned technology base programs. Section 5.3 next describes how programs tools and infrastructure are factored into the planning process. Section 5.4 then summarizes the specific plans for initiating the overall program.

The material in this chapter is intended to provide an overview sufficient to enable readers to interpret detailed Strategic Computing planning documents, such as the timelines in the appendices, and to have well-formed intuitions concerning the overall methods and plans of the program.

On a first reading, those who are interested in specific dimensions of the program might read the Chapter 5 section of interest (for example the applications) and skim the rest. Alternatively, the logic of the plan can be sampled by following the details of one chain of activities through the material (without reading all sections in detail). For example, one could follow the requirements for vision produced and passed from the Autonomous Vehicle (Section 5.1.1; App. I.1) through to the section on vision subsystems (Section 5.2.1.1; App. II.1.1), and then to the sections on System Architecture, and Infrastructure. In that way the interplay between the applications and technology base can be closely examined.

5.1 Applications Experiments and Demonstrations

This section describes a set of specific military applications that exploit new generation computing technology. Included is an autonomous vehicle application that will rely heavily on vision and expert system technology as enabling technologies for this application. A pilot's associate application is then described that exploits speech recognition and expert systems. Finally, a battle management system for a carrier battle group is described that exploits expert systems technology and that will eventually exploit very high performance knowledge processing systems.

These specific examples are included in the Strategic Computing program based on a weighted consideration of the following factors:

- The application must effectively employ the new technology to provide a major increase in defense capability in light of realistic scenarios of combat situations that might occur at the future time when the new systems can be procured and deployed.
- The application must provide an effective "pull" on the new generation technology. It must demand an aggressive but feasible level of functional capability from one or more of the intelligent functions at appropriate points in the timeline.
- Development of the application must lead to new engineering know-how in artificial intelligence software areas, such as planning and reasoning, learning, navigation, knowledge base management, and so on.
- The application must test the efficacy of the new technology at a realistic quantitative scale of performance demands. In this way we seek to ensure against unexpected quantitative changes in system performance as a result of scaling up from models and laboratory experiments to real systems.
- The application must provide an effective experimental "test-bed" for evolving and demonstrating the function(s). Stability over time, access, and visibility are thus important factors.

- o The application must effectively leverage program resources. Thus an important factor is the extent to which an existing military program provides a base of capital resources and experienced personnel into which the new generation technology can be experimentally introduced (versus this program having to provide such non-computing resources).
- o It is important to choose a mix of applications that are jointly supportive of and involve all three Services, and which are appropriately executed through each Service. Only in this way can we develop the base for extension of this technology into a wide range of military systems.
- o Finally, an important selection factor is the potential provided by the specific application for effecting the transfer into the services of the new machine intelligence technology.

The choices have been initially made on this basis, but it is recognized that further planning and the evolving technology development may lead to a change in the choice of specific application demonstrations. It might, for example, prove preferable to pursue an autonomous underwater vehicle rather than a land vehicle, and a battle management system for land combat might prove more appropriate than that for the Naval application. A panel of the Defense Science Board has been convened to make recommendations on how to best exploit machine intelligence technology within DoD, and that panel will be providing information and advice for this program. Consequently, we anticipate that some of the specifics may change over time, within the framework that is described.

An abbreviated description of each currently planned application is given in this section. Planning timelines are included in Appendix I that illustrate the detailed interactions of these applications with ongoing military programs and with the emerging new generation technology base.

5.1.1 Autonomous Vehicles

Autonomous systems, as used herein, are true robotic devices: they are able to sense and interpret their environment, to plan and reason using sensed and other data, to initiate actions to be taken, and to communicate with humans or other systems. Examples of autonomous systems include certain "smart" munitions, cruise missiles, various types of vehicles possessing an autonomous navigation capability, and a wide variety of mobile and fixed robotic systems for material handling, manufacturing, and other applications. Some of these systems exist today with operationally useful levels of capability. Others, such as completely autonomous air, land and undersea vehicles, and systems possessing more adaptive, predatory forms of terminal homing, require the kinds of significant developments anticipated in the Strategic Computing program to fully realize their potential. These developments will both enable qualitatively different kinds of autonomous behavior in new systems and effect dramatic quantitative improvements in the operational capabilities of existing systems.

Autonomous vehicles, like other autonomous systems, are characterized by their ability to accept high-level goal statements or task descriptions. For an autonomous vehicle system one set of goal statements will define a navigation task, another will be specific to its mission, for example reconnaissance. The navigation task will usually be described both in terms of a specific destination for the vehicle and through constraints which limit the number of possible paths or routes the vehicle might use in traversing from one point to another.

Autonomous land vehicle systems, as an example of a class of autonomous vehicles, could support such missions as deep-penetration reconnaissance, rear area re-supply, ammunition handling, and weapons delivery. As an example, imagine a reconnaissance vehicle that could navigate up to 50 km cross-country from one designated position to another. It would be capable of planning an initial route from digital terrain data, updating its plan based on information derived from its sensors, resolving ambiguities between sensed and pre-stored terrain data, and incorporating landmark prediction and identification as a navigation means. Using advanced image understanding technology, the reconnaissance payload would perform image

segmentation and other basic scene processing upon arrival in a designated area, identify target objects, and report its findings and interpretations.

To develop an autonomous land vehicle with the capabilities described requires an expert system for navigation and a vision system. The expert navigation system must plan routes using digital terrain and environmental data, devise strategies for avoiding unanticipated obstacles, estimate the vehicles' position from landmark and other data, update the on-board digital terrain data base, generate moment-to-moment steering and speed commands, and monitor vehicle performance and on-board systems. All these functions must be accomplished in real-time to near-real-time while the vehicle is moving at speeds up to 60 km/hr. Scaling up from laboratory experiments indicates that such an expert system demonstration would require on the order of 6,500 rules firing at a rate of 7,000 rules/second. Current systems contain fewer rules, 2,000 on average, and fire at a rate of 50-100 rules/second. ("Rules and firings" are terms used in expert systems. "Rules" represent the codification of an expert system process and a "firing" indicates the examination, interpretation, and response to one rule in a particular context. In current systems, the firing of one rule can require the execution of tens of thousands of instructions, and as contexts become more complex the number of instructions for rule-firing increases.)

The vision system must take in data from imaging sensors and interpret these data in real-time to produce a symbolic description of the vehicle's environment. It must recognize roads and road boundaries, select, locate, and dimension fixed and moving obstacles in the roadway; detect, locate, and classify objects in open or forrested terrain, locate, and identify man-made and natural landmarks, and produce thematic maps of the local environment, while moving at speeds up to 60 km/hr. Scaling up computing capabilities used in laboratory vision experiments suggests an aggregate computing requirement of 10-100 BIPs (billion equivalent von-Neumann instructions per second) to accomplish the above tasks. This compares with capabilities, for example, of 30-40 MIPs (million instructions per second) in today's most powerful von-Neumann type computers.

Of equal importance with the required computing capabilities outlined above, is the weight, space, and power required by the computing systems. For a land reconnaissance vehicle, for example, the computers should occupy no more than 6-15 ft^3, should weigh less than 200-500 lbs., and should consume less than 1 kw of power including environmental support. The requirements represent at least 1-4 orders of magnitude reduction in weight, space, and power over today's computing systems. For certain space, air, and sea vehicles, the constraints and requirements will be even higher and will include the capability to operate in high-radiation environments.

5.1.2 Pilot's Associate

Pilots in combat are regularly overwhelmed by the quantity of incoming data and communications on which they must base life or death decisions. They can be equally overwhelmed by the dozens of switches, buttons, and knobs that cover their control handles demanding precise activation. While each of the aircraft's hundreds of components serve legitimate purposes, the technologies which created them have far outpaced our skill at intelligently interfacing the pilot to them.

This mismatch seems to be characteristic of many human controlled, complex, dynamic military systems. Further, it applies to single operator as well as multiple operator situations where crew communication and coordination are essential for survival. It is this type of common military problem that pulls intelligent computing technology into the realm of creating the "personal associate."

The personal associate is viewed as an ensemble of expert knowledge based systems and natural interface mechanisms that operate in real-time. In its simplest form the personal associate performs a set of routine tasks and, when prearranged, initiates actions on its own. In this way it frees the operator from routine overhead chores so he can attend to more critical tasks.

In its advanced form, the personal associate performs a set of tasks which are difficult or impossible for the operator altogether, such as the early detection and diagnosis of the subtle patterns of an impending malfunction. In this way the associate enables completely new capabilities and sophistication.

We have chosen to illustrate this concept by developing a personal associate within the context of the combat pilot. Called the Pilot's Associate, it is an intelligent system that assists the pilot in the air as well as on the ground, not replacing but complementing the pilot by off-loading lower-level chores and performing special functions so the pilot may focus his intellectual resources on tactical and strategic objectives.

The associate is personal to a specific pilot in that it is trained by that pilot to respond in certain ways and perform particular

functions. For example, it might be instructed to automatically reconfigure the aircraft to a specific control sensitivity preferred by the pilot should the wing be damaged during combat. It also has a wealth of general knowledge about the aircraft, the environment, and friendly and hostile forces. It will have instruction on advanced tactics from more experienced pilots and up-to-date intelligence information on enemy tactics to aid the less experienced pilot on his first day of combat. These knowledge bases will be designed for easy updating to keep pace with rapidly changing tactical events. Certain classes of newly "learned" knowledge will be automatically exchanged among pilot's associates.

The approach for this application is to evolve an increasingly complex pilot's associate in increments that represent key program decision milestones. The development will be continually evaluated in full mission research simulators with representative combat pilots, eventually to be moved onboard existing research aircraft for evaluation. The three thrusts central to this development are: the interface to the pilot, the knowledge bases to support the interface, and integration and interpretation processors that connect these.

The interface is based upon natural communication using advances in speech recognition (here developed for the noisy, stressful cockpit environment), speech output (particularly machine speech that can assume different speaker types and styles), and graphic or pictorial presentation of complicated information.

Knowledge bases will be developed that will be significantly larger than any previously attempted. For example, the simple monitoring of the basic flight systems (power, electrical, and hydraulic) could take several thousand rules. These will have to be processed at rates perhaps 100 times faster than the current technology allows. The knowledge bases that will be developed are:

- The aircraft/pilot
- Communication
- The mission
- Tactics and strategy
- Geography
- Enemy defense
- Enemy aircraft
- Navigation aids
- Friendly forces

The processes that integrate and interpret the demands from the interface with the contents of the knowledge bases include functions that tie flight events to the mission plan prepared earlier, change environmental and threat situations, coordinate with other pilot's associates in the air and on the ground, continually change situational data bases for local battles as they develop, and so forth.

The demand for realtime processing eventually in small, rugged packages for onboard installation characterizes the "pull" that this application puts on the Strategic Computing program. The knowledge gained will directly complement important Service research programs such as the USAF Cockpit Automation Technology effort.

5.1.3 Battle Management

Management of large scale enterprises is characterized by decision making under uncertainty. The system must alert the decision maker to the existence of an incipient problem, must generate potential responses to the problem in the form of decision options, must evaluate these options in the face of uncertainty about the outcome arising from any specific option and with respect to often conflicting goals, must execute the preferred option, and monitor its execution, iterating on the above process as circumstances dictate. No examples exist today of systems which directly address each of the above steps. While many individual information processing systems, such as the World Wide Military Command and Control System (WWMCCS) and various intelligence systems, furnish data to the decision maker that support such functions as alerting and option generating, the fact remains that no systems exist which directly aid such cognitive processes as option generation, uncertainty assessment and multi-attribute value reconciliation. These are knowledge intensive and the development of aids in these and other critical areas will consequently require the kinds of expert system and natural language developments anticipated from the Strategic Computing Program.

A battle management system (BMS), as an example of a system to aid in the management of a large enterprise, would interact with the user at a high level through speech and natural language. It would be capable of comprehending uncertain data to produce forecasts of likely events, drawing on previous human and machine experience to generate potential courses of action, evaluating these options and explaining the supporting rationale for the evaluations to the decision maker, developing a plan for implementing the option selected by the decision maker, disseminating this plan to those concerned, and reporting progress to the decision maker during the execution phase.

For example, a Battle Management System for a Carrier Battle Group would be integrated into the Composite Warfare Commander (CWC) battle group defense system. It would display a detailed picture of the battle area, including enemy order of battle (surface, air, sub-surface), own

force disposition, electronic warfare environment, strike plan, weather forecast, and other factors developed from an analysis of all available data. It would generate hypotheses describing possible enemy intent, prioritize these according to their induced likelihood, and explain the reasons for the prioritization. Drawing upon previous experience, together with knowledge of own force and enemy capabilities, it would generate potential courses of action, use an ultra-rapid rule-based simulation to project and explain a likely outcome for each course of action, and evaluate and explain the relative attractiveness of each outcome considering such criteria as protection of own forces, inflicting damage on the enemy and the rules of engagement. Once the commander selects a course of action, the BMS would prepare and disseminate the operation plan (OPLAN), and compare the effects of option execution with those developed through the simulation both as a check on progress and as a means of identifying the need to replan. At the conclusion of every phase of the engagement, the BMS would modify its expert system in the light of empirical results.

The Naval Carrier Battle Group Battle Management System (see Chart I.3, Appendix I for details) builds upon experience and developments in the existing DARPA/Navy program to utilize expert systems and display technology on the Carrier USS Carl Vinson, and can exploit potential associated opportunities of the CINCPACFLT command center ashore.

To realize the capabilities described above will require the development of a number of expert systems and a natural language interface. The expert systems, for the demonstration BMS will make inferences about enemy and own force air order-of-battle which explicitly include uncertainty, generate strike options, carry out simulations for evaluating these strike options, generate the OPLAN, and produce explanations. It is estimated that in the aggregate the above functions define a distributed expert system requiring some 20,000 rules and processing speeds of 10 BIPs. The natural language system alone will require a processing speed of about 1 BIP.

Space-based signal processing requirements for surveillance and communications will require low power, very high speed, radiation hardened, integrated circuits based on gallium-arsenide technology. These circuits will operate at speeds of at least 200 megahertz, with tens of milliwatts of power required for a typical 16 kilobit memory, in radiation up to 5 x 10^7 rads.

While the preceeding text described a Battle Management System for a Carrier Battle Group, it must be emphasized that the impact of the technology base required for this development extends substantially beyond the scope of this specific application. For example, many of the hardware and software developments would support, with different data, Army tactical battle management at the corps, division and battalion level, logistics management, and missile defense.

53-831 O -86 - 17

5.2 The New Computing Technology

This section describes specific technology areas in the new generation technology base. In Section 5.2.1 we discuss three of the integrated "intelligent" functions: vision, speech, and natural language understanding, along with expert system technology as a means of implementation. These are areas where considerable progress has already been made, and these functions will be inserted into applications experiments early in the program.

For these areas it will be possible to codify laboratory knowledge in order to produce generic software systems that will be substantially independent of particular applications. A variety of other software areas, such as planning and reasoning, are not now as well developed. At the beginning of the program they will be pursued in the context of particular military applications in order to produce engineering know-how that can be extended to a broad range of problems. We anticipate that at a later time in the program some of these other software areas will be sufficiently well understood that they, too, can be developed to provide application independent software "packages."

A short description is given of each area, and a timeline for each area is included in Appendix II. These timelines can be cross-compared with the timelines for the functions' applications (Appendix I) and architectural implementation (Appendix II). It is important to note that a number of other intelligent functions will be competitively inserted into the program at later times, as basic research matures and as the applications environments provide opportunities for their experimental development and demonstration.

Section 5.2.2 describes the technology area of hardware/software system architecture and design. This is the key area of the structural design of machines and software to implement the intelligent functions. The parameters of specific designs are set where appropriate by specific applications experiments in which the machines will be used. This section

suggests the manner in which the applications and the intelligent functions' requirements will "pull" the architecture and design of new generation machines. The reader can cross-compare the summary timeline for Section 5.2.2 (Appendix II) with the timelines for sections in 5.1 (Appendix I) and section 5.2.1 (Appendix II).

Section 5.2.3 describes the area of microelectronics. The Strategic Computing Program will place great emphasis on the effective exploitation of state of the art microelectronics (see also section 5.3) in order to meet the key constraints on power, weight, volume, and performance required by the selected applications. The development of the GaAs pilot lines is specifically included within the Strategic Computing program. The remainder of the supporting microelectronics technology is ongoing in the basic DARPA program, or under development by industry, and will contribute directly as results become available.

5.2.1 Integrated Intelligent Functions

5.2.1.1 Vision. Computer vision, also called image understanding, is the information-processing task of comprehending a scene from its projected image. It differs from related disciplines such as pattern recognition and image processing in that the process of image understanding builds a description not only of the image data itself, but also of the actual scene which is depicted. Image understanding requires knowledge about the task world, as well as sophisticated image-processing techniques.

DARPA has carried on a basic research program in computer vision for some years. The technology has matured to the point where it can now be exploited in meaningful ways. Since the autonomous vehicle application described previously stresses the technology development to a significant extent, it will serve as the initial driver of technology research. In order to meet objectives of the vehicle application, generic recognition capability will be required for both vehicle navigation and for reconnaissance. A vision subsystem will have to provide for the recognition and identification of obstacles that might deter local navigation and also for landmarks that can be used to fix vehicle position in the global navigational sense. The vision component must also be able to recognize targets and understand, at least from the standpoint of threat evaluation, what is happening to objects of interest from scene to scene. To achieve these capabilities, specific advances will have to be made in both vision software and also in the hardware that will run the necessary computer programs (see Appendix II, Chart II.1.1).

There currently exist software algorithms that perform object recognition in highly specific task domains but techniques will have to be developed that generalize this capability. Furthermore, the recognition process will have to be robust enough to permit recognition in the face of occlusion, shadows, and differing orientations. The key to achieving this capability is significant advances in high-level modeling and use for knowledge-based recognition techniques. These concerns will receive strong emphasis in the early stages of the project. There will also be efforts to implement discrimination capabilities to differentiate objects of interest,

e.g., discerning obstacles as opposed to landmarks or targets. As the system evolves, it should also develop the capacity to detect moving objects within its range of vision, understand that they are moving, and comprehend the relations of their movement to other objects in the scene.

Recent progress in developing vision for navigation has been severely constrained by lack of adequate computing hardware. Not only are the machines which are now being used too large to be carried by the experimental vehicles, but current machines are far too slow to execute the vision algorithms in real-time. For example, in an experimental university research program, a "corridor rover" applied a vision subsystem to navigate itself down corridors that had various obstacles in its path. The scene is re-analyzed after the cart has moved 1 meter. The current algorithms require 15 minutes of compute time for each meter moved. If the vehicle were moved at a walking pace, the computing requirements would be about 3 orders of magnitude greater. Future applications will have more complex scenes, be required to move faster, and also require the performance of various tasks en-route.

It is estimated that 1 trillion von Neumann equivalent computer operations per second are required to perform the vehicle vision task at a level that will satisfy the autonomous vehicle project's long-range objectives. At best, current machines of reasonable cost achieve processing rates below 100 million operations per second. The required factor of 10^6 improvement in speed will have to be achieved through VLSI implementation of massively parallel architectures. In order to make use of these architectures, parallel algorithms will have to be developed. Therefore, part of the early research efforts will also concentrate on the development of suitable parallel algorithms. It is felt that low level vision processes can be exploited in a more straightforward fashion because of the inherent parallel nature of images and the local operations that are performed. Thus, initial emphasis will concentrate on algorithms at this level. As a better understanding of the problems is gained, the parallel programming efforts will evolve to embrace the higher-level vision processes.

The most significant technology that will result from this effort is a generic scene understanding capability. This technology will be exportable to a wide range of military applications, including cruise missile en-route navigation and terminal homing, as well as a wide variety of fire-and-forget weaponry.

Functional Objectives for Vision Subsystems

FY 86	Model and Recognize Simple Terrain with Crude Objects
FY 88	Recognize and Match Landmarks with Maps in Simple Terrain
FY 90	Recognize and Match Landmarks and Obstacles in Complex Terrain Using Rich Object Descriptions
FY 92	Perform Reconnaissance in a Dynamically Changing Environment

5.2.1.2 Speech Recognition and Production. The program goal for speech subsystems is to enable real-time speech input to computers and the generation of meaningful acoustic output. Past efforts in speech understanding have been limited by both inadequate processing capabilities and by an inadequate understanding of the acoustic phonetics of speech. On-going basic research programs in speech are addressing a number of the basic issues. This program will capitalize on the results of this basic research.

The capabilities of a speech subsystem vary along several dimensions that include:

- isolated word recognition to continuous speech
- speaker dependent to speaker independent
- quiet environments to noisy, stressful environments
- small vocabularies in limited context to vocabularies having 10,000 or more words

This program is concentrating on developing speech recognition and generation high-performance capabilities for two generic types of applications: One in a high-noise, high-stress environment, where a limited vocabulary can be useful, such as in the fighter cockpit, and another in a moderate-noise environment where a very large vocabulary is

required, such as in a battle management system. The timeline for this program is shown in Appendix II, Chart II.1.2, including specific milestones. The technology is applicable to many other tasks; the applications cited here provide a focus for the research and insure that these specific applications will be supported with speech.

For the cockpit application, the major challenge will be to develop speech recognition algorithms which can operate in a fighter aircraft environment. This includes noise levels up to 115 dB, acceleration to several g's, voice distortions due to the helmet and facemask, and the changing voice characteristics under the stresses of combat. The initial computational requirements are estimated to be 40 MIPS to demonstrate speech recognition in the cockpit, counting both the signal processing and recognition functions. Furthermore, this hardware must be sufficiently compact so as not to exceed the restricted space and power that is available in a fighter aircraft.

The initial set of tasks focus on speaker-dependent isolated word recognition in a noisy environment. The specific use of speech recognition in the cockpit needs to be studied in detail to understand which tasks should be performed by voice, how voice will impact other systems, what vocabulary is needed, etc. Speaker dependent algorithms for recognizing words in a noisy environment will be developed initially, and will later be extended to speaker independent algorithms. A prototype architecture for performing the real-time recognition tasks will be developed and used to evaluate algorithms in a simulated cockpit environment. This initial architecture would be composed of off-the-shelf hardware and would not be suitable for flight. Compact hardware will be developed, including custom hardware for performing compute intensive functions such as template matching.

Support of spoken natural language input and output for a battle management system will require real-time continuous speech recognition and generation of very large vocabularies of 1,000 to 10,000 words with natural syntax and semantics, in a relatively benign acoustic environment. Techniques will need to be developed for the acquisition and representation of knowledge of speech variability due to alternate pronunciations, context in

continuous speech, and different speakers. Efficient parallel search algorithms and hardware, combined with techniques for focusing attention on key works, will be developed for dealing with large vocabularies. Automated techniques will be developed for acquiring the acoustic, syntactic, and semantic knowledge to switch among multiple task domains. Advanced acoustic-phonetic algorithms will be needed to distinguish among similar words in large vocabularies. Integration of the speech system with the natural language system will be required to perform the overall battle management task.

Increasing speech capabilities will be developed over time, with an initial goal of 1,000-word speaker-adaptive system, and an ultimate goal of a 10,000-word speaker-independent system. We estimate that the computational requirements of the latter system will be on the order of 20 BIPS.

Functional Objectives for Speech Subsystems

FY 86 Recognition of Words from a 100-Word Vocabulary for a Given Speaker under Severe Noise and Moderate Stress Conditions

FY 88 Recognition of Sentences from a 1,000-Word Vocabulary with Moderate Grammatical Constraints in a Speaker Adaptive Mode under Low Noise and Stress Conditions

FY 89 Recognition of Connected Speech, Independent of Speakers from a 200-Word Vocabulary with Strict Grammatical Constraints under Severe Noise and High Stress Conditions

FY 92 Recognition of Sentences, Independent of Speakers, from a 10,000-Word Vocabulary with Natural Grammar under Moderate Noise and Low Stress Conditions

5.2.1.3 Natural Language Understanding. The most common way for people to communicate is by expressing themselves in a natural language such as English. If we can produce computer programs that can deal with a substantial subset of English meaning, we can make headway on several fronts. In the first place, we can provide natural language interfaces so that tactical experts can be closely coupled with supporting databases and

automated expert systems. Such interfaces would accept data inputs, commands, and queries in natural language and could furnish responses either in natural language or in the form of easily understandable text and tables. We can also develop systems that understand streams of text to achieve automatic input of information transmitted in that form.

Natural language research has matured to the point where it is finding application as a man-machine interface in various commercial equipments. However, its application to operational military environments is still limited by the lack of sufficient computing capacity, an inadequate understanding of semantics and discourse context, inadequate vocabularies, and the conceptually challenging and time consuming problem of introducing sufficient knowledge and semantics into the system. Ongoing basic research programs will address some of these issues and feed into this program but additional intensive efforts are needed to achieve the technology level necessary for meeting the requirements of the Battle Management application described elsewhere in this plan.

The technology subprogram in natural language has the overall objective of achieving an automated understanding and generation capability that can be used in a variety of applications. We will undertake research that supports this objective by focusing on the technology needed to fulfill the specific natural language requirements of the Battle Management problem. This approach will not only support the implementation of a Battle Management system, but progress made in this area will also be applicable to a wide class of similar problems. Meeting the requirements will entail the development of a highly intelligent natural language interface between the user and the machine. In addition, a text processing component will be developed that can classify text by its context, determine and store the key events, and retrieve the relevant information by contextual reference with an accuracy of no less than 95%. The timeline for this subprogram is shown in Appendix II, chart II.1.3.

In order to achieve the desired capability of the natural language front end, it will be necessary to make significant advances in three specific areas. First of all, natural language understanding programs must have a much greater comprehension of the context of the ongoing discourse

between the user and the machine. This will significantly reduce the amount of dialogue that has to take place by instilling the capability within the machine to anticipate requirements of the user. Secondly, a much more sophisticated level of natural language response on the part of the machine is required so that information can be presented in the most meaningful way to the user. Thirdly, an interactive facility for the acquisition of knowledge has to be developed. This is driven by the time-consuming requirements of incorporating new linguistic and semantic knowledge in the system. In the area of text understanding, advances must be made in the area of cognitive memory modeling and text comprehension.

In order to develop the capability we envision, several milestone systems will be built. The first of these will integrate and slightly extend existing natural-language interface techniques. There will then be a dual effort, one aimed at text processing and the other at interactive dialogue systems. Each of these efforts will result in specialized intermediate milestone systems. Finally, these streams will be joined together to achieve the full functional capability necessary to suport the Battle Management application.

<u>Functional Objectives for Natural Language Subsystems</u>

FY 86 Natural-language interfaces with some understanding of commands, data inputs, and queries (e.g., interface to a database and a threat-assessment expert system)

FY 88 Domain-specfic text understanding (e.g., understand paragraph-length intelligence material relating to air threat)

FY 90 Interactive planning assistant which carries on task-oriented conversation with the user

FY 93 Interactive, multi-user acquisition, analysis, and explanation system which provides planning support and substantive understanding of streams of textual information

The tasks described above will require substantially larger vocabularies than are currently available and significant gains in processing

power in order to accomplish understanding and response in real time. It is estimated that vocabularies of 15,000 words and processing speeds of 1 billion operations per second will be needed to achieve this goal. In addition, to be useful for practical applications, this power must come in compact dimensions. These constraints will generally necessitate the utilization of massively parallel VLSI computational devices. Such an architecture will in turn demand the reformulation and development of parallel algorithms for natural language understanding.

5.2.1.4 Expert System Technology. Expert System technology has matured to become a highly exploitable application area of the science of artificial intelligence. It is characterized by the explicit use of specific domain knowledge (usually gleaned from human experts) to develop computer systems that can solve complex, real-world problems of military, scientific, engineering, medical and management specialists.

Examples of successful applications include programs to perform electronic warfare signal analysis, medical diagnosis, geological evaluation of designated sites, oil well dipmeter analysis, maintenance of locomotives, and carrier air operations. It is a technology that is most appropriate for command and control operations, situation assessment, and high-level planning. Thus it will play a vital role in the military applications examples described elsewhere in this plan.

Expert system technology has evolved to a point where a variety of general purpose inferencing and reasoning systems are available. These systems can be augmented with specific domain knowledge to prepare them for particular applications. Currently, the most time consuming portion of the process of constructing an expert system is the articulation of knowledge by the expert and its satisfactory formulation in a suitable knowledge representation language for mechanization by computer. Thus, the plan for expert systems technology (see Appendix II, Chart II.1.4) places heavy emphasis on knowledge acquisition and representation.

There are many opportunities for dramatic advances in the technology. These include advances in explanation and presentation capability, improved ability to handle uncertain and missing knowledge and data, more flexible control mechanisms, expansion of knowledge capacity and extent, enhanced inference capability (in terms of speed, flexibility, and power), development of inter-system cooperation, and improvement of software support tools. Intensive development attention devoted to these issues can be expected to lead to important applications of expert systems in complex military environments.

The Strategic Computing expert system technology effort will exploit these opportunities by generating and extending AI techniques, by improving software support tools, and by using specialized symbolic computational hardware. Work in representation will build toward a capability for large (30,000 rule) knowledge bases. Inference techniques will be extended to handle these knowledge bases even when they contain uncertain knowledge and must operate on errorful and incomplete data. Explanation and presentation systems, ultimately using a 10,000 word speech understanding system, will allow verbal inputs from (and discussions with) the user about the systems' assessments, recommendations and plans. The knowledge acquisition work will focus on developing facilities for automated input of domain knowledge directly from experts, text, and data. Software support efforts will lead to a progression of increasingly powerful expert system workstations to be used in developing the needed technology.

The achievement of these complex capabilities will severly tax computational resources so that significant gains in processing power will be required to perform in real time or in simulations at faster than real time. It is estimated that hybrid expert system architectural configurations will be required that can accommodate 30,000 rules and perform at a capacity of 12,000 rules per real-time second at rates up to 5 times real time. Due to compact size and cost constraints, it is anticipated that this architecture will be realized through VLSI devices incorporating massive parallelism, active semantic memories, and specialized inference

mechanisms. Such configurations will require significant efforts to develop the algorithms required for parallel execution. It should be noted that the rules per second quantifications are subject to many factors, and are for comparison purposes only. Rules applied in applications late in the program will be more complex than present ones, and their contexts for firing will be vastly more complex than those common in present-day expert systems.

The results of this effort will specifically support the goals of the three example military applications. However, the resulting technology will be substantially generic in nature so that it will significantly advance expert systems capabilities and support a wide-range of applications for both the Government and industry.

5.2.2 Hardware/Software System Architecture and Design

Most of today's computers are still single-processor von Neumann machines, and the few efforts to build commercial multiprocessor systems have yielded systems containing only a few processors (generally less than 10). The underlying electronic circuit technology is advancing at a rate that will provide a speed improvement factor of only 20 to 30 percent per year, at most, for such machines. For future computer systems to have substantially greater power, they must rely heavily on parallelism. While many ideas have been developed for algorithms, languages, and system software for high performance parallel machines, practical experience with actual experimental parallel systems is still very limited, and must be greatly expanded.

Greater computing power can also be achieved through specialiation of machines to particular computing functions. Such specialized machines exhibit exceptional performance, but only on the class of problems for which they were designed. Parallelism is itself a form of specialization of a machine to a class of problems. For example, array processors will out-perform a comparably priced general purpose computer by factors of 10 to 100 on linear algebra, finite element analysis, and similar problems. Future high performance systems for applications such as the control of autonomous vehicles must support a diverse and demanding set of functions with high reliability. Such systems will be composed of a variety of modules configured to perform these many specialized functions efficiently, in parallel, and with redundancy appropriate to the application. For example, the control of autonomous vehicles may employ modules specialized to signal processing to handle the image processing at the lowest level, modules specialized to pattern matching to handle the scene analysis, and other modules to handle cognitive functions, control, and communications. This integration of diverse machines into complete systems depends on standardization of hardware, software, and network interfaces.

Computer architecture is concerned with the structures by which memories, processing nodes and peripherals are interconnected; the computational capabilities of the processing nodes; and the software which is required to exploit the hardware. Ideas have been proposed for machines

which are interconnected in a variety of ways, and given descriptive names such as Boolean n-cubes, trees, perfect shuffles and meshes. Processor nodes have been proposed that are designed for floating point operations, search operations, logic operations, etc., and language and operating system concepts have been proposed for exploiting parallelism. It is from this collection of ideas that specific architectures have been proposed, and in some cases simulated or constructed on a very small scale.

To understand the capabilities and limitations of a proposed architecture, a prototype of the machine must be simulated or built, software must be developed, and the system evaluated on a class of problems for which the machine was designed. The role of software cannot be overemphasized. Existing languages are generally not applicable for highly parallel architectures. Special compilers are needed, as are debugging tools and tools to measure the performance of the resulting system. In evaluating a new architecture, it is more important to initially understand the applicability of the architecture to an important class of problems than to strive for high performance in a prototype implementation. Thus, to know that a 100 processor system gives a 50 fold increase over a single node of that system is more important than knowing the maximum instruction rate that can be executed or knowing the exact instruction rate achieved with prototype hardware. Once a prototype machine has been demonstrated to be promising, higher performance versions can be built by using faster components and by scaling the entire system to have more processors.

This program will develop and evaluate new architectures in 3 broad areas: signal processing, symbolic processing and multi-function machines. These classes of machines are described below, along with the development plans for each class. In general, several prototype systems will be developed in the early phase of the program. An evaluation phase will permit different architectural approaches to be compared. We will select from the different prototypes those which are most successful and which will be continued to develop high performance versions.

A timeline for the development of these computer architectures is given in Appendix II, Chart II.2.2.

5.2.2.1 <u>Signal Processing</u>. An important class of applications known as signal processing involves taking real-time data from a sensor and performing a series of operations on each data element. These operations might involve transformations such as an FFT, correlations, filtering, etc. and are dominated by performing multiplications and additions. High data rates are common, and computation rates in excess of 1 billion operations/second are needed. Military applications of such signal processing include processing data from radar, sonar, infrared sensors, images and speech.

The exploitation of parallelism in signal processing will be based on the use of computational arrays such as systolic arrays, in which many simple, highly regular processing elements "pump" data from cell to cell in a "wave-like" motion to perform the successive operations on each element of data. An architecture based on this concept will be developed, with the goal of building a system capable of executing 1 billion or more operations/second by 1986. Other concepts that exploit signal processing data regularity will also be investigated. By the end of a decade, the goal is to develop a system capable of 1 trillion operations/second.

The software support and programming languages for the signal processing system will be developed in parallel with the hardware. Most of the initial programming for the prototype system will be done at the microcode level. The requirements for operating system, programming languages and programming environments will be developed as experience is gained using the prototype systems.

5.2.2.2 <u>Symbolic Processing</u>. Symbolic processing deals with non-numeric objects, relationships between these objects, and the ability to infer or deduce new information with the aid of programs which "reason." Examples of symbolic computation include searching and comparing complex structures (e.g., partial pattern matching). Applications which make extensive use of symbolic computing include vision systems which can tell what is in a scene, natural language systems which can "understand" the meaning of a sentence in English, speech understanding systems which can recognize spoken words, and planning systems which can provide intelligent advice to a decision maker. Most programs which perform symbolic

processing are now written in the language called LISP. Special machines, called LISP machines, are now available commercially and offer computing rates in excess of one MIP. Further development of these conventional uniprocessor Lisp machines will take place under the technology infrastructure portion of the program. An ultimate performance improvement of about 50 times the current level can be achieved with these conventional techniques and the use of advanced technology.

Current applications in areas such as vision now require about three orders of magnitude more processing than is now available. As future algorithms and applications are developed, even more computing power will be necessary.

The symbolic processors of the future may well be a collection of special components which are interconnected via a general purpose host computer or by high speed networks. Based on software systems which have been developed for applications in vision, natural language, expert systems, and speech, several of these components have been identified. As much as four orders of magnitude speedup may be available by taking advantage of the parallelism in some of these specific areas. Some of the components include the following:

- A semantic memory subsystem -- used to represent knowledge relating concepts to other concepts in natural language, speech understanding, and planning domains.
- A signal to symbol transducer -- used to make the initial step in extracting meaning from low level signal processing computations (e.g., phonetic classification, or object identification from boundary information).
- A production rule subsystem -- a system that combines knowledge and procedures for problem solving. A system now aboard the carrier Carl Vinson uses this approach.
- A fusion subsystem -- a method for permitting multiple sources of information to share their knowledge. It is used to "fuse" information in tasks such as battle management.

- o An inferencing subsystem -- a system that uses first order formal logic to perform reasoning and theorem proving.
- o A search subsystem -- a mechanism that explores numerous hypotheses, pruning these intelligently to determine likely candidates for further symbolic processing.

The program will consist of three phases. Phase I concentrates on architecture design, simulation, algorithm analysis, and benchmark development for promising architectural ideas such as those described above. It will also include the development and initial evaluation of the unique integrated components necessary for the implementation of these architectures. The design of concurrent LISP-like languages for programming these machines will also be addressed.

Existing high-performance scientific computers such as the Cray-1, CDC 205, Denelcor HEP, and the S-1 will be benchmarked using a portable LISP computer to determine their relative abilities to handle symbolic computation.

Phase II will engineer full scale prototype versions of selected architectures, supporting these hardware developments with extensive diagnostic and compilation tools. The goal of this phase is implementation of a specific target problem on each of the selected architectures for benchmarking purposes.

Phase III will integrate developments of the signal processing, symbolic, and multi-function development efforts into a composite system capable of addressing a significant problem domain. Such a system for the control of an autonomous vehicle, for example, might include a high performance vision processing front-end based on the computational array technology, a signal-to-symbol transformer for classifying objects, a fusion subsystem for integrating information from multiple sources, an inferencing engine for reasoning and top-level control, and a multi-function processor for controlling the manipulator effectors. This phase will also pursue higher performance versions of selected machines.

5.2.2.3 Multi-Function Machines. A multi-function machine is capable of executing a wider range of different types of computations than

the more specialized machines described above, but at possibly lower performance in the specialized machine's application domain. These multi-function machines achieve high performance with parallelism. We aim to develop machines of this class having 1000 processors. The processing elements in a multi-function machine would typically be general purpose processors or computers. These elements communicate either through shared storage or networks with such interconnection strategies as rings, trees, Boolean n-cubes, perfect shuffle networks, lattices, or meshes.

On the order of 6 to 8 prototype multi-function systems will be developed, based on custom VLSI chips, commercial microprocessor chips, or commercial processors. These systems will be benchmarked to determine how different hardware architectures and programming strategies scale in performance. Subsequently, 2 or 3 such systems will be selected in this evaluation process for continued development for advanced technology versions and production quality software.

Central to this program is the development of programming models and methods which will permit the convenient development of new classes of algorithms which will contain very high levels of concurrency. The way in which concurrency manifests itself in program structures can be viewed as resulting from the linguistic control method of the programming language in which the program is written. Examples of control models which will be investigated are the control-driven, data-driven, and demand-driven styles. Control-driven concurrent programming models are already evolving from existing programming languages. Examples are concurrent PASCAL, parallel LISP, etc. In this model, program actions are sequenced by explicit control mechanisms such as CALL, JUMP, or PARBEGIN. In the data-driven model, program actions are driven into activity by the arrival of the requisite operand set. The advantage of this style is that concurrency can often be specified implicitly. The demand-driven model is based on the propagation of demands for results to invoke actions. This style has been successfully employed for parallel evaluation of LISP code. In this scheme, concurrent demands are propagated for argument evaluation of LISP functions. It is likely that new or possibly composite models such as

concurrent object oriented programming will surface, but it is also likely that advances in each area will provide highly-concurrent program-based solutions for many application areas.

An important part of this project will be the implementation of new concurrent programming languages which exploit these models. The language development will need to be coupled with programming environment tools and compatible hardware and operating system software. This development will provide the necessary computational tools to support application studies aimed at the creation of highly parallel application programs which can take advantage of the large levels of concurrency provided by multi-function machine prototypes. The long term goal of this research is ultra speed, cost effective demonstrations of important application areas such as data base access, system simulation, and physical modelling.

Given a particular instance of machine and a particular parallel program, the remaining issue is how the program should be mapped onto the physical resources in order to permit efficient exploitation of concurrency. This resource allocation problem is one of the key technical issues addressed by this program. There are two styles usually employed in the solution of this problem: static allocation and dynamic allocation.

In a static mapping strategy, the concurrency structure of the program is evaluated with respect to the topology of the physical machine. The compiler can then create specific load modules for the physical nodes of the target machine. This static method is simpler than the dynamic method but needs to be developed for each of the architectures which are being pursued. If the number of components is very large, then it is likely that component failures will occur. With the static allocation mechanism, it will be necessary to recompile the program for the current machine configuration.

In a dynamic allocation strategy, it is still important for the compiler to do some of the allocation task collection but the output of the compiler is not in the form of specific load modules. Dynamic strategies allow the loader to define the final physical target of a compiled module

based on hardware availability. Another extension to the dynamic strategy is to additionally move tasks around to balance system load. An important byproduct of this program will be the development and implementation of both static and dynamic allocation strategies but it is expected that acceptable static allocation methods will precede the more sophisticated dynamic strategies.

5.2.3 Supporting Microelectronics Technology

Computing technology relies heavily on microelectronics in order to achieve systems capabilities while meeting critical constraints on such factors as size, weight, power dissipation and operating environments. Microelectronics provides computing systems with required integration complexity, switching speed, switching energy and tolerance to hostile environments. In the case of military systems, special emphasis must be given to survival in radiation environments. Microelectronic packaging and interconnect technologies provide additional important support in meeting system constraints.

This program will place strong emphasis on the effective exploitation of such microelectronic technology. A key concept that will be used to exploit state-of-the-art microelectronics is to dramatically reduce the usual long-time delays between basic research innovations in fabrication and packaging technology and their subsequent exploitation by designers. This will be done by creating a pilot line(s) for the particular technology and at the same time creating the associated designer-to-implementation system-to-foundry interfaces (design rules, process test inserts, design examples, design libraries, implementation system protocols, etc.). Once a new technology has been demonstrated as feasible and stable in pilot-line form, it may then be selected for inclusion in program infrastructure. [The reader should compare the microelectronics timeline (5.2.3) with that for infrastructure (5.3). (See Appendix II.)]

Silicon Technology. Silicon technology will be the mainstay of this program because of its maturity and its accessibility through existing infrastructure. Early versions of the proposed subsystems prototypes will use the 3×10^{11} gate hertz/cm^2 technologies made possible by VLSI/VHSIC. More advanced technologies such as the VHSIC Phase II will also be utilized as they become available. For subsystems and/or systems that require even greater throughput this program will competitively purchase wafer level integration technology from emerging sources. This will result in the gain of at least another order of magnitude in the computational throughput of a monolithic chip and an equally significant reduction in power consumption

for a given operation by diminishing the number of required off-chip drives.

Even such gains will not fulfill the ultimate weight, volume, and speed requirements of such systems as certain autonomous vehicles that will require better than 10^{10} operations per second and 10^{11} bits of memory in less than a few cubic ft using no more than a kilowatt of power. For these requirements to be met, new fabrication technology must be developed yielding devices an order of magnitude smaller than those produced today. Ultimately, techniques now in basic research phases such as ion-beam processing technology, laser processing and x-ray lithography may be combined with silicon molecular beam epitaxy into a pilot line system capable of growing multiple semi-conductor and insulator layers, adding localized ion doping, etching via holes and depositing interconnect metal. If successful in moving from basic research, such efforts could eventually reduce from month to days the time required to fabricate prototype custom circuits of high complexity.

GaAs Pilot Lines. Survivable, space based electronics will require the two orders of magnitude increased total dose radiation tolerance that is inherent to GaAs based microelectronics technologies. The establishment of pilot lines, running at a throughput of at least 100 wafers/week, will place, for the first time, the production rigor on the fabrication of GaAs integrated circuits necessary to achieve acceptable yields and make GaAs circuits affordable to the military. The GaAs pilot lines will be producing low-power, radiation-hard memory and logic chips as fundamental building blocks for radiation hardened systems. Communication and surveillance systems that can survive in a strategic conflict are important components of a space-based battle management system. In addition to the primary advantage of high radiation tolerance, GaAs based microelectronics will also produce circuits with larger operating temperature range, both lower and higher than silicon, and faster on-chip switching speed at a given power level.

Memory Technology. Rapid-access, low-power memory subsystems that can be operated in the field and powered from conventional sources are

needed by many applications. Today's largest disk storage systems contain on the order of a gigabyte of memory but are too large and power-consuming for use in the field. Progress must be made in both size and power reduction. Systems needs for as large as 100 gigabyte memories with rapid access are envisioned for autonomous systems. The program will capitalize on progress in industry and in other basic research programs.

High Performance Technology. The need to increase system computational speeds may be met using fabrication technology that can tailor materials properties by creating artificial compounds and super-lattices of differing materials (for example by using Molecular Beam Epitaxy (MBE) technology). Successful pilot-lining of MBE would contribute to conventional microelectronics, microelectronics with opto electronic I/O, and eventually to massively parallel computations using optical computing elements. When available optoelectronic interconnect technology will allow the number of cables and the power dissipation in large multiprocessor systems to be reduced dramatically.

As Molecular Beam Epitaxy (MBE) systems advance into a practical production tool, heterostructure devices can be fabricated to produce high frequency devices. Such a development will reduce transmit-receive satellite systems presently requiring 6 ft. dishes to possibly hand-held devices by utilizing the 94 GHz atmospheric window. Such systems will contribute in a revolutionary manner to size, weight and cost of battle management communications subsystems.

Advanced computing subsystems-on-a-chip will require both high speed and high pin-count packages. The VHSIC program is developing 250 pin packages but these are only suitable up to a 40 megahertz clock rate. This program will initiate development and/or compete the selection of large pin-count packages, including those that employ microwave signal propagation principles. Longer term efforts directed towards achieving (optoelectronic) packages required to handle broad band operation, from d.c. to multi-gigahertz, with up to 200 signal lines in addition to power and ground leads will be factored into the program where appropriate.

5.3 Computing Technology Infrastructure

In order to effectively support and coordinate the activities of the large number of people and organizations in this program, we will focus our attention early in the program on the provision of adequate infrastructure for the enterprise. It is intended that this be accomplished in a manner that rapidly disseminates the technology, not only across the participants, but across US industry.

There are three phases of activity in this part of the program. Beginning in the first year, major emphasis will be placed on the consolidation of state-of-the-art computing technology to enable rapid capitalization and maximum resource sharing. A second phase is designed to take advantage of early products of the program to enhance overall capabilities. A final phase is intended to bring about a transition of the activities in the infrastructure to make them self-supporting. These three phases result in a cost profile which is initially high, but is reduced over the life of the program as costs are gradually borne by recipients of the technology. High initial investment in computing equipment, services and training also leverages the most critical resource -- trained personnel.

The infrastructure is categorized by specific activities to be performed. The most immediate need is for availability of the products of computing technology so as to bootstrap the development process. We will provide common symbolic processing equipment to the selected participants in the first two years and will supplement this with more advanced equipment as it is developed in the program. Common access to high performance networks will be provided to facilitate communication between sites and shared use of computing resources.

Next, a set of activities address common access to services and tools that are the means of designing and building new computers. These include rapid-prototyping implementation services providing foundry access to VLSI/VHSIC and GaAs fabrication lines as well as access to higher-level system implementation services. Computers are to be used extensively in the design and analysis of new systems and these hardware and software

tools will be shared between sites by exploiting the common hardware configurations, programming languages, and network communication facilities.

Finally, the most important activities in the infrastructure accelerate the rate of progress. These activities appear as items integral to the products and services just described as well as specific activities in their own right.

The following examples are representative of the methods sought to achieve this acceleration:

- Use of state-of-the-art computing technology to develop new computing technology.
- Shared access to capital intensive manufacturing facilities.
- Improved productivity through use of advanced design methods and system interoperability kits.
- Rapid turnaround implementation services.

In addition to these, specific activities encourage collaboration between researchers through the development of interoperability standards. Strong interaction with the university community is coupled with the use of the technology in the form of embedded instruction to accelerate the training and development of personnel.

The result is a powerful expansion of the traditional concept of infrastructure. The program will produce not only an advanced technology base in the form of facilities, equipment, institutions, and knowledge, but also the methods for using it and accelerating its growth.

5.3.1 Capital Equipment

Hardware and software will be developed early in the program to enable widespread use of advanced symbolic computers and communication systems in both laboratory and embedded applications.

5.3.1.1 LISP Machines. A small number (25-40 per year) of LISP machines will be acquired during the first years of the program for use by contractors in the conduct of research and applications development. In parallel two new classes of LISP machines will be developed by industrial manufacturers. One will be 10x faster than current machines and the second will be a low power, compact version for use in applications experiments,

field trials, and demonstrations. This equipment will be supplied to contractors beginning in FY87.

5.3.1.2 Research Machines. As new machine capabilities are developed and demonstrated in the program, other defense projects will be able to benefit from direct access to them. We plan to develop some of these machines and supply them with the necessary software and intelligent subsystems for use in follow-on R&D in support of the military demonstrations. Other systems will be available via a network. Industrial production of these machines will be sought where appropriate.

5.3.1.3 Communication Networks. No element of the infrastructure is as important as the need for widening the network connection among the various participants of the program. Beyond the obvious advantages of sharing resources and facilities, the network is unparalleled as a means of promoting synergy between researchers located at different sites. We plan to work with other agencies (such as DOE, NASA and NSF) in developing a common plan for leased wideband communication facilities to be made available by the common carriers.

5.3.2 Services

The physical construction of complex computing equipment is a difficult and time-consuming task, even when all the essential design details are understood. A set of services will be put in place that simplify this process, reduce cost, and provide rapid turnaround.

5.3.2.1 Integrated Circuit Implementation Service. Silicon VLSI/VHSIC and GaAs fabrication lines will be made available as foundries for use by selected Defense contractors. We plan to work with the vendors to develop standard design rules for this technology and provide access via network connections. This will extend the method already in use for 3-5 micron NMOS and CMOS of providing direct access from the designer's system over the network to the foundry service. This service will be expanded to provide access to advanced microelectronic technology as it is developed under 5.2.3.

5.3.2.2 Rapid Machine Prototyping. A service will be established to allow the rapid implementation of full scale systems with the

goal of enabling the assembly of complete multiprocessor computer systems from initial designs in a period of a few months. This service will provide rapid turnaround services for printed circuit boards, hybrid fabrication, system packaging, power, cooling, assembly and testing.

5.3.2.3 System Interoperability Kits. Sources will be solicited for the design and manufacture of "system kits" intended to facilitate interoperability and experimentation in new computer architectures. These standardized hardware/software environments will provide the physical means of easily integrating and assembling systems into predesigned modules using design frames for embedding unique custom designs as part of these systems.

5.3.3 Tools for an Integrated System Development Environment

An advanced system development environment will be constructed as a framework for consolidation and integration of the design and performance analysis tools that are produced by this program. This environment will set the standards for tool development and facilitate the sharing of the products of this research between sites. A major benefit of this common system development environment when coupled with rapid prototyping is that it allows hardware decisions to be deferred and an optimal balance of hardware and software achieved.

5.3.3.1 Functional and Physical Design Aids. This new generation of computers will be developed using new high level tools that are built upon state-of-the-art research in VLSI design. These tools will be extended upward to enable system level design, assembly, and test in a rapid system prototyping environment. It is here that the use of computing technology as a tool to create new computing technology is most obvious. In the functional design of a new machine architecture, its performance can be evaluated through emulation. We expect to use dedicated hardware emulation machines to assess a number of architectures for which construction will be difficult or costly. Likewise, advanced hardware and software approaches to physical design aids will enable more rapid and robust system design.

5.3.3.2 Software and Systems. An integrated rapid software and systems prototyping capability is needed to support the development and application of multi-processor systems. This capability will be developed by building upon advanced software and systems development environments such as ADA and LISP and extending them to support multi-processor targets. The major problems that need to be solved to effectively apply these architectures and achieve the required performance and resource allocation for processor, memory, communication, and mass storage. In addition, the application system developers need support for using the new architectures in terms of the virtual machine interfaces that will be developed to manage resource allocation.

The Software and Systems activities produce the most generic software to support the application specific software. This includes programming languages, system software, and design and performance analysis tools for multiprocessor targets. As the technology matures, resource allocation will become more automatic and higher level design environments for multiprocessor architectures will be developed.

5.3.4 Standards

To integrate hardware and software to perform basic system functions, and then to integrate those functions into systems will require interoperability research. A key ingredient will be the set of protocols that allow interaction between modules. It should be possible to access information in a knowledge base from a speech understanding system or to make available vision or natural language to a navigation system. Outputs from any of these should be available to AI based simulation and display systems.

We envision developing system interoperability protocols to the point where couplings of hardware, software, and peripheral devices may be selected and configured readily. This will include capabilities for speech input, vision, graphics and a host of intelligent system tools including an expert system and a LISP machine.

5.4 Program Planning

So far in this chapter, we have presented the key concepts of the Strategic Computing plan by showing how example activities proceed under the plan. We now step back and summarize the overall logic of the plan, list the compartments and activities to be planned, and discuss the detailed tactics to be used to initiate the Strategic Computing program.

As first discussed in Chapter 4 and clarified by example in Chapter 5, the top-level logic of the plan centers on the interactions of selected military applications of intelligent computing with the evolving base of technology that provides the intelligent computing. In particular,

- o Applications drive requirements of intelligent functions.
- o Intelligent functions drive requirements of system architectures.
- o System architectures drive requirements of microelectronics and infrastructure

In order to achieve the Program's goals, we must create on the order of a dozen different, modularly composible intelligent functional capabilities. Each one, such as vision subsystems, requires the generation of a "technical community" responsible for evolving that technology.

However, since these functions are broadly applicable to many applications, it is likely that a modest number (perhaps a half-dozen) well-selected applications will be sufficient to "drive" the whole set of intelligent functions. Each of these applications similarly requires the generation of a technical community, prime contractor, or center of excellence responsible for its evolution. A range of hardware/software architectures must also be created, systems must be implemented as microelectronics, and adequate infrastructure must be provided to support the entire enterprise.

Later in Chapter 6 and in Appendix IV, we provide a detailed work-breakdown structure that compartments all these program activities for planning and budgeting purposes. We now turn to the plans for initiating the program.

We are initially concerned with the development of appropriate military applications that will effectively pull the technology base. A set of three applications have been selected for initial inclusion in the program. Based on the results of a Defense Science Board task force study (see Chapter 6), and a series of competitive evaluations of the most impactful applications, we plan to augment and refine the list to a final set during the first several years of the program, using selection criteria cited in Section 5.1.

At the beginning of the program we will initiate work in the four areas of intelligent functions (vision, speech, natural language, expert systems) that can be exploited in the near-term, and drive these technologies using requirements set by the selected applications. At later times we will initiate activities (as basic research matures) in the other area of intelligent capabilities.

Activities will begin in system architecture on two fronts. The first will be development of systems aimed at supporting the near-term intelligent functions for selected applications (an example would be a computational array processor to support vision technology). Next, we plan competition among several large symbolic processor architectures that will be prototyped for later evaluation and selection. Such processors will be essential during later stages of the program. Additional specialized system architectures will be selected, later in the program timeline at points where they are required by applications.

Certain microelectronics technology will be developed in pilot-line form early in the program (for example GaAs), in order to position the technology for support of later program requirements.

A very key portion of the plan for program initiation is the early development and deployment of program infrastructure (see Section 5.3 for details). Research machines, network communications, implementation services, etc., must be in place to enable program progress. A set of protocols and interoperability standards must be created to insure later modular compatibility among Strategic Computing technology components. As we will

see in Chapter 6, this means that spending on infrastructure is a moderately high fraction of program spending in the first two years (although it rapidly peaks and levels off).

Finally, appropriate program management support tools must be brought on-line early in the program to insure orderly, planned, managed progress toward program goals and objectives.

When studying the Program Timelines in the Appendices, note that the planning framework is not a closed system "PERT" chart, but instead is open to accommodate available opportunities and competively selected technologies during appropriate time-windows. Thus it is the generation of the technology envelope that is planned, charted, and guided to achieve program goals, rather than the generation of a specific computer or specific technology module.

CHAPTER 6
PROGRAM MANAGEMENT

The management of the Strategic Computing Program will be carried out by the Defense Advanced Research Projects Agency. This chapter describes DARPA's approach to management of the program. Because of the importance, size, complexity, and pace of this program a number of issues will be addressed.

Program Coordination. The importance of the Strategic Computing Program to the national interest requires coordination with many different organizations involved in related technologies.

Within DoD, DARPA will coordinate closely with USDRE and the military services. Preliminary discussions have been held with representatives of all three Services, and all have expressed strong interest in close cooperation with DARPA on this program.

An agreement for exchange of information has been reached among OSTP, DOE, NSF, NASA, DOC, and DOD representatives at a spring meeting of the Federal Coordination Committee on Science Engineering and Technology sponsored by OSTP. This agreement called for a series of meetings specifically organized to exchange information in high speed computing technology. The first of these meetings was held in June 1983, with DoD chairing the meeting. Further meetings will be scheduled at regular intervals until the end of the program or until otherwise mutually agreed.

A panel of the Defense Science Board (DSB) headed by Professor Joshua Lederberg, President of Rockefeller University, has also been convened to make recommendations to the Under Secretary on how best to use the new-generation machine intelligence technology within DoD.

Program Management. The Strategic Computing Program will be managed within DARPA. The Strategic Computing Program Manager will be assigned to the Information Processing Techniques Office (IPTO), the lead Office. However, significant responsibilities are allocated to other offices, especially in the areas of microelectronics and applications.

53-831 O - 86 - 18

It is DARPA's objective to maintain a dynamic R&D environment for this project and to manage the delicate balance between the technology-base development and the experimentation in military applications.

The number of active working relationships may be very large because other offices within DARPA, USDRE, universities, the Services, and industry will all be involved. Since these relationships must be maintained to ensure integrated planning and execution, the program will use advisory panels to reach these groups. One of these will be a senior review group to provide advice as the program progresses. It will consist of representatives from the three services, OSD, other governmental organizations, and major industrial organizations and universities. In this way the Program will capture the best creative ideas of government, universities, and industry while continually involving the ultimate user community. Thus group will meet quarterly with the DARPA management involved in the program.

Similarly, other panels or working groups will be constituted to provide communications and advice in specific areas and to keep other groups abreast of progress in Strategic Computing.

The Strategic Computing Program planners will continue refinement and adaptation of the program plan over time to reflect the current state of funding and development. Efforts will be undertaken to maintain program documentation, provide technical evaluation of progress, respond to congressional (and other) inquiries, develop internal reporting and control systems, support program reviews, maintain technical libraries, and disseminate information in the form of technical abstracts and progress reports, as required.

Communication is a critical element of program management because many of the important contributors will be widely dispersed. Special and unique arrangements will be considered to establish an effective research community by leveraging existing computer tools and communications systems. Electronic mail and electronic bulletin boards are the simplest examples. More advanced approaches to be considered include the provision of remote electronic views of the unfolding project planning timeline to give

feedback on performance to community members. This is an innovative approach that derives from successful experiences in DARPA program activities in VLSI system design. In this way we will build a planning and management infrastructure that enables new "intellectual entrepreneurs" to easily identify ways they can contribute to the program to create new elements of intelligent computing technology and its military applications.

Program Costs and Work Breakdown Structure. The overall scale of budget requirements is determined by the number and type of new technologies to be jointly introduced, and by past experience in the field in generating equivalent technological communities or centers of expertise. In this case we require the creation of approximately ten new "computing-technology communities" and another five to ten "applications communities." The scale of size of these efforts must, if past experience is any guide, be at least on the order of a small research center (>100 professionals) each composed of two or three research laboratories, that operate over approximately an eight to ten year time span. This scale is consistent with past requirements for the generation of new computing technology such as timesharing, computer-networking, personal computing, etc. This logic provides a top-down scaling of the enterprise as requiring about 5 to 10 applications plus 10 technology communities of at least 100 persons each, and thus requires something approaching 150 million dollars per year for a several year period around the peak of the program.

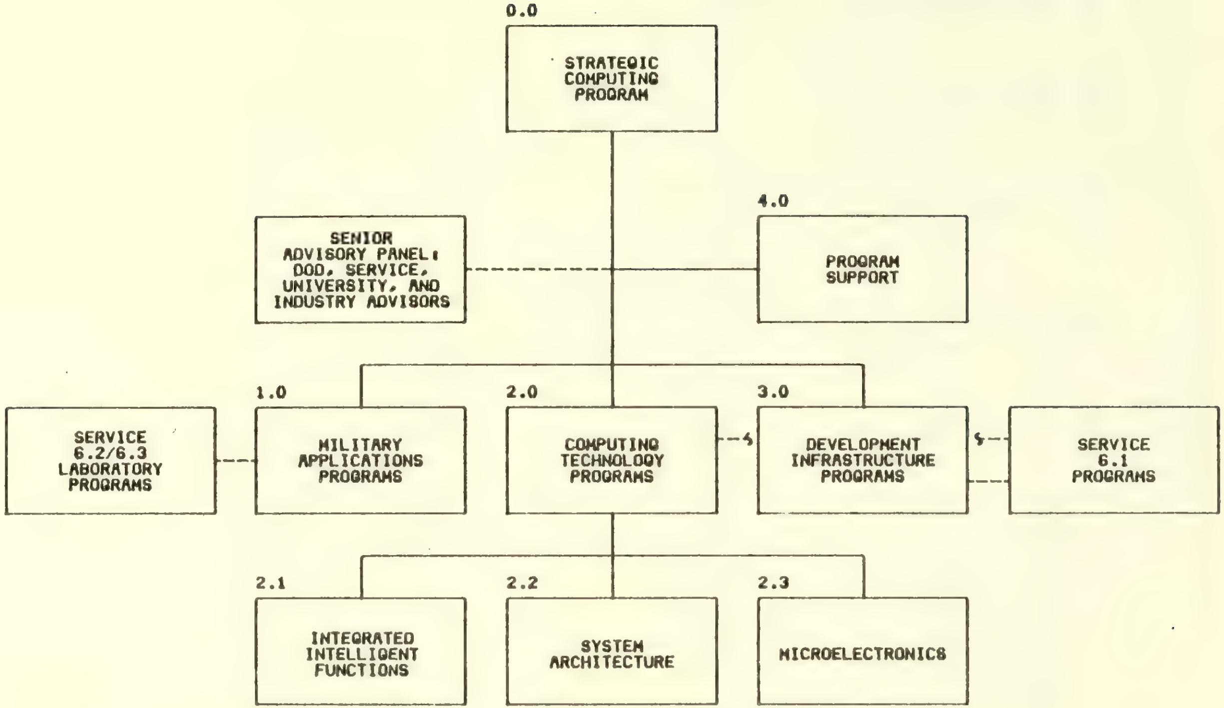
This estimate is consistent with a bottom-up budget estimate based on the detailed unfolding of actual and projected activities for the initial period of the plan where specific projection can be readily made.

A program management and work breakdown structure for the program is diagrammed in Figure 6.1. Further details of the work breakdown structure are tabulated in Appendix IV. Figure 6.2 shows the annual cost for the Strategic Computing Program aggregated to the program level. Program costs for the first five years of the program are estimated to be approximately 600 million dollars.

The logic of the sequencing of activities is reflected in the breakdown of spending in the three major categories. Spending on tools and infrastructure is relatively high in the first two years, peaking early in the program. Technology base activity and spending then rises fast, and will likely peak in FY 87-88. Applications activity and spending expand moderately at first, then rapidly in the late 80s, peaking near the end of the program. The entire program will peak about the end of the decade, declining rapidly thereafter as program goals are achieved.

Acquisition Strategy. The basic acquisition policy is that military applications will be carried out by industry drawing upon results of research carried out in the universities. The computer architectures will be developed primarily in joint projects between universities and industry. Most of the hardware and software for intelligent subsystems will be competed. There will be a selection of ideas on especially difficult topics from a set of several dozen leading contenders. The most advanced artificial intelligence ideas that seem ripe for developing will be exploited with heavy university involvement. For these, expert judgment from leading participants in the field will be sought and directed selection will result. Construction and access to computing technology infrastructure will be competed.

The contract personnel responsible for accomplishing the goals of this program will be largely drawn from industry and will consist primarily of engineers and systems designers. By contracting with industry we will transfer to that community computer science research results that have been developed in universities, largely with DARPA funding; we will ease the transition of the newly developed systems into corporate product lines; we will avoid a dangerous depletion of the university computer science community, with the inevitable slow-down in research and education. The magnitude of this national effort could represent a very large perturbation to the university community, but is a small percentage of the industrial engineering and system-building base.



W6196AR3284

FIGURE 6.1 DARPA PROGRAM MANAGEMENT AND WORK BREAKDOWN STRUCTURE FOR STRATEGIC COMPUTING

STRATEGIC COMPUTING COST SUMMARY IN $M

	FY84	FY85	FY86	FY87*	FY88*
Total Military Applications	6	15	27	TBD	TBD
Total Technology Base	26	50	83	TBD	TBD
Total Infrastructure	16	27	36	TBD	TBD
Total Program Support	2	3	4	TBD	TBD
TOTAL	50	95	150	TBD	TBD

* Out-year funding levels to be determined by program progress.

Figure 6.2

While most of the basic technology development in this program will be unclassified, the emphasis on industrial efforts will provide a significant control of the leakage of information outside of the US industrial base.

Technology Transfer. We intend a significant effort toward technology transfer of results of this program into the military services. This effort will include: (a) use of Service Agents and Service COTRs; (b) a process of cost-sharing with the Services in the development of military applications; (c) the inclusion of technology base results from this program in Service Programs and Testbeds, and (d) training of Service personnel by involvement in technology base developments.

Equally important is technology transfer to industry, both to build up a base of engineers and system builders familiar with computer science and machine intelligence technology now limited to university laboratories, and to facilitate incorporation of the new technology into corporate product lines. To this end we will make full use of regulations for Government procurement involving protection of proprietary information and trade secrets, patent rights, and licensing and royalty arrangements.

Evaluation. Each of the sections of this program will have a detailed evaluation plan. Specifically:

(1) Each of the microelectronics developments will be proposed to particular performance specifications, e.g., radiation hardness for GaAs, and the final deliverable will be evaluated against those specifications, driven by requirements.

(2) Early in the program, benchmark programs will be developed for evaluation of the competitive machine architectures. For example, different signal processor designs will be benchmarked with programs drawn from radar or sonar analysis, or some other chosen signal analysis task; different symbolic processors will be benchmarked through tasks such as evaluation of a particular production rule set, or searching through a particular semantic set, such as one used for natural language understanding; and,

more general purpose processor designs will be benchmarked with applications that might involve war gaming or simulation.

(3) Performance requirements for the integrated intelligent functions, i.e., vision, speech, natural language and expert systems will be defined by the requirements of the three (or more) chosen military application areas, and evaluations will be performed toward those specifications.

(4) Finally, the military applications developed will be evaluated using the same methods and criteria currently used by the Services. This will simplify comparison. For example, the evaluation of the efficacy of the pilot's associate will be measured in combat performance - - with and without - - on instrumented combat flight ranges.

CHAPTER 7
CONCLUSIONS

We now have a plan for action as we cross the threshold into a new generation of computing. It is a plan for creating a large array of machine intelligence technology that can be scaled and mixed in countless ways for diverse applications.

We have a plan for "pulling" the technology-generation process by creating carefully selected technology interactions with challenging military applications. These applications also provide the experimental test beds for refining the new technology and for demonstrating the feasibility of particular intelligent computing capabilities.

The timely, successful generation and application of intelligent computing technology will have profound effects. If the technology is widely dispersed in applications throughout our society, Americans will have a significantly improved capability to handle complex tasks and to codify, mechanize, and propagate their knowledge. The new technology will improve the capability of our industrial, military and political leaders to tap the nation's pool of knowledge and effectively manage large enterprises, even in times of great stress and change.

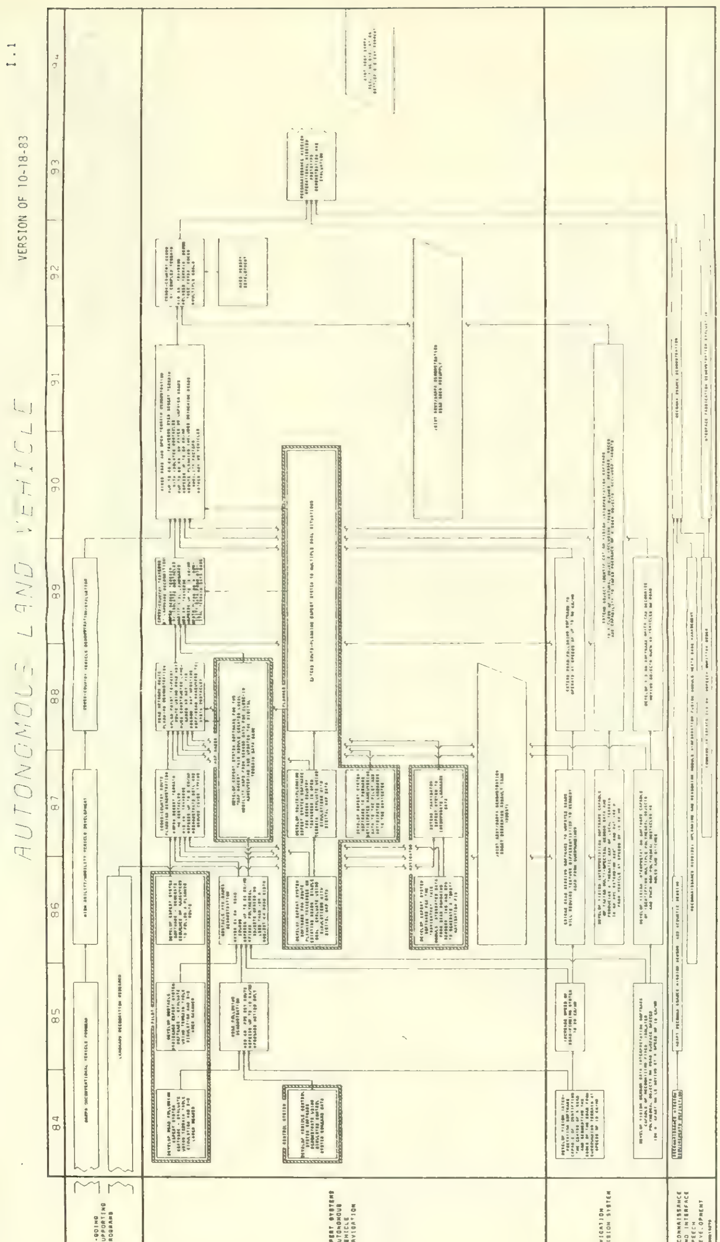
Successful achievement of the objectives of the Strategic Computing initiative will lead to deployment of a new generation of military systems containing machine intelligence technology. These systems will provide the United States with important new methods of defense against massed forces in the future - methods that can raise the threshold and diminish the likelihood of major conflict.

There are difficult challenges to overcome in order to realize the goals of a national program of such scope and complexity. However, we believe that the goals are achievable under the logic and methods of this plan, and if we seize the moment and undertake this initiative, the Strategic Computing Program will yield a substantial return on invested resources in terms of increased national security and economic strength.

APPENDIX I

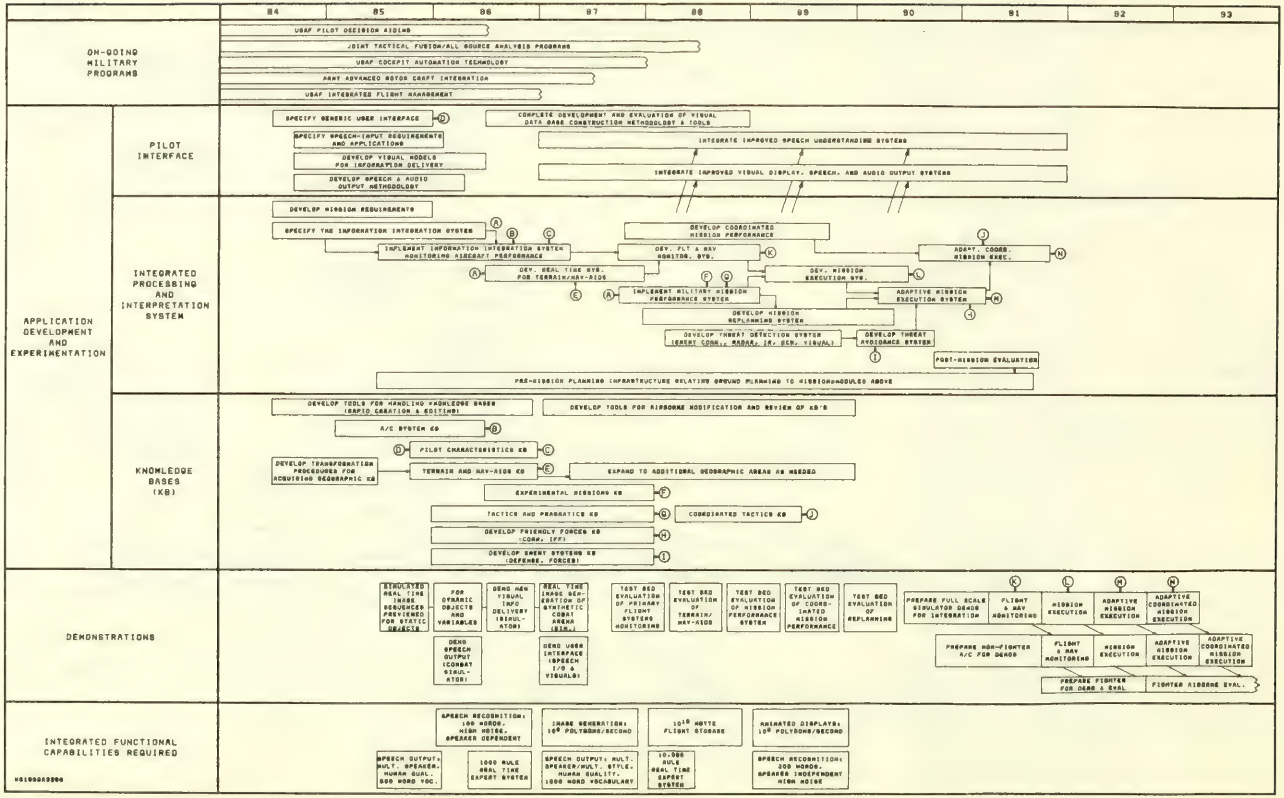
MILITARY APPLICATIONS PLANS

This appendix contains a set of planning timelines for the Strategic Computing military application examples discussed in Chapter 5 (Section 5.1). These timelines illustrate how the applications interact with ongoing military programs, and how the applications pass functional requirements to the emerging new generation computing technology (see also Appendix II).


AUTONOMOUS LAND VEHICLE
VERSION OF 10-18-83
I.1
84
85
86
87
88
89
90
91
92
93
94
ON-GOING SUPPORTING PROGRAMS
EXPERT SYSTEMS AUTONOMOUS VEHICLE NAVIGATION
NAVIGATION VISION SYSTEM
RECONNAISSANCE AND INTERFACE SPEECH DEVELOPMENT

PILOT'S ASSOCIATE

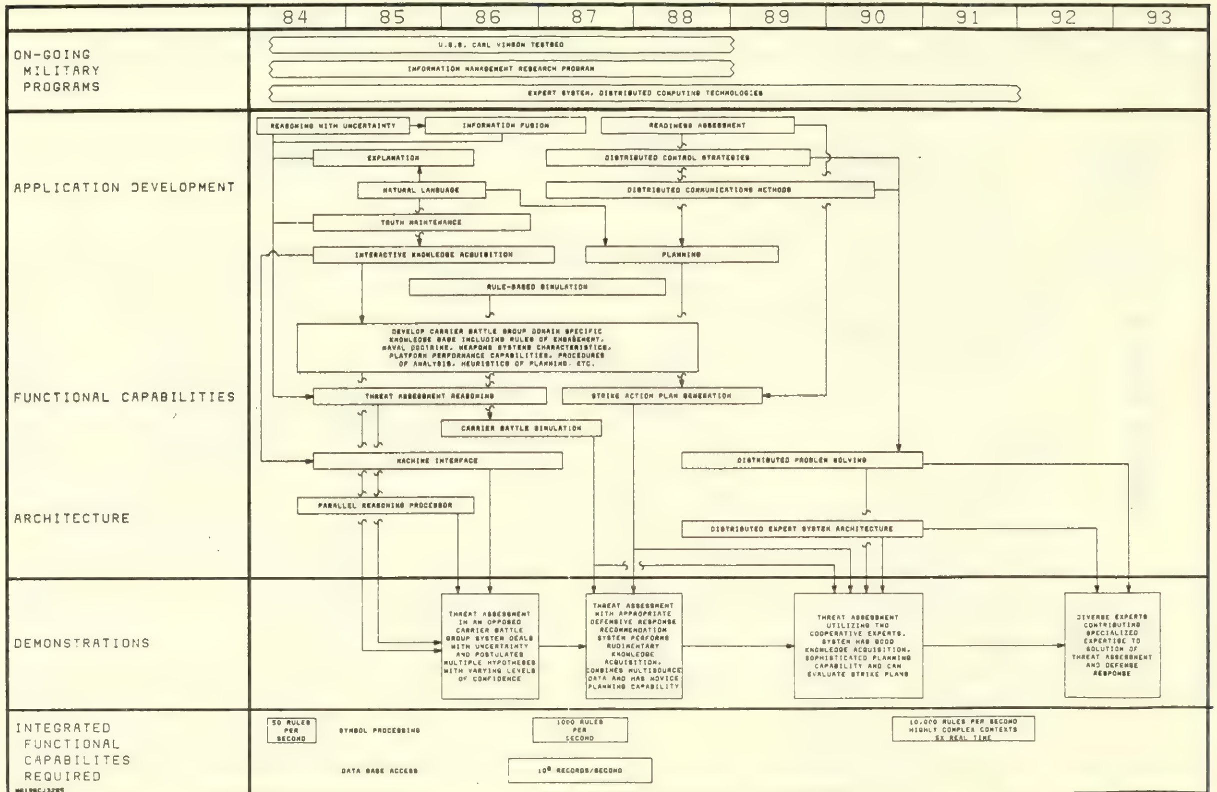
VERSION OF 10-18-83 1.2



VERSION OF 10-18-83

CARRIER BATTLE GROUP BATTLE MANAGEMENT SYSTEM

I.3



APPENDIX II

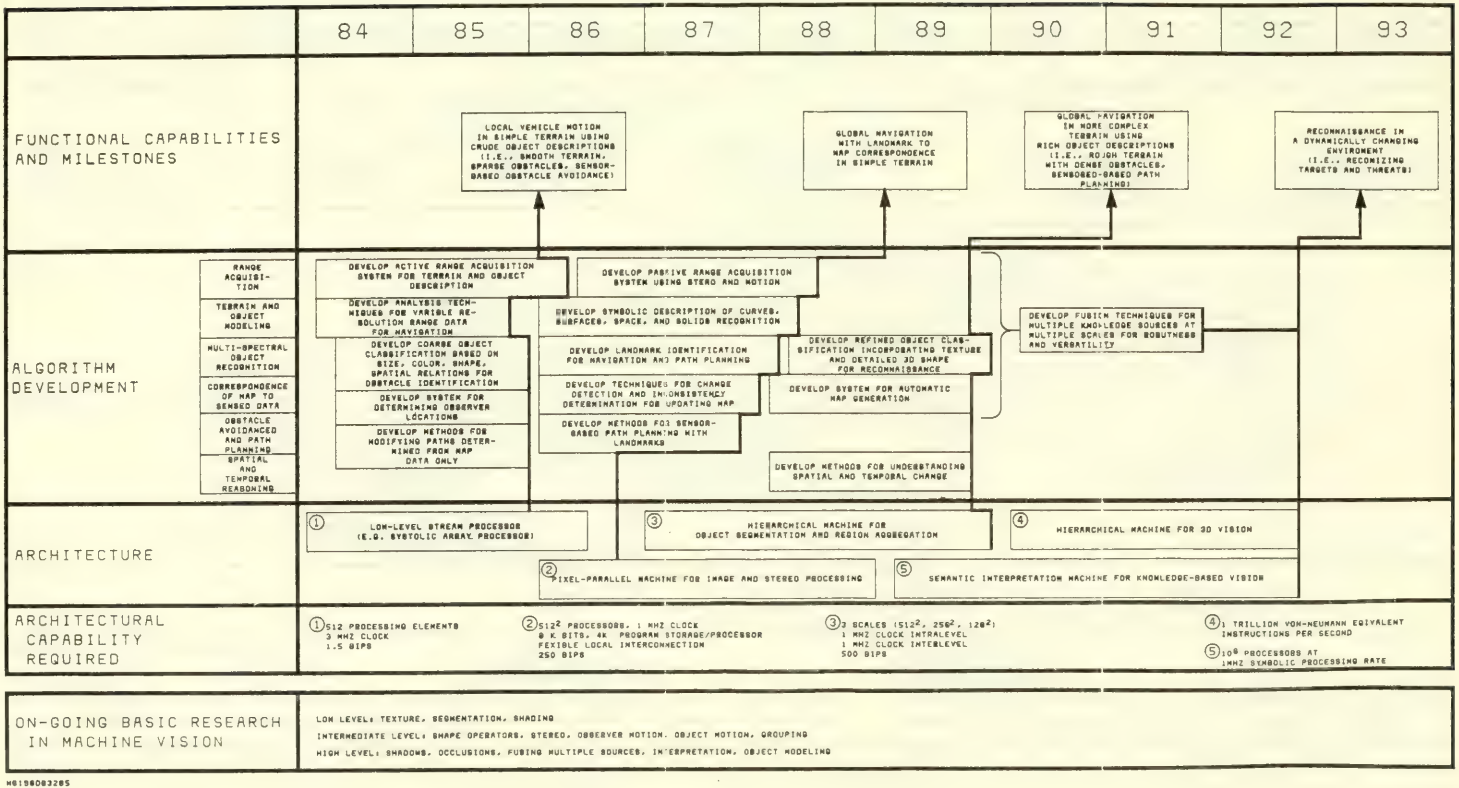
COMPUTING TECHNOLOGY PLANS

This appendix contains a set of planning timelines for the Strategic Computing technology base examples discussed in Chapter 5 (Section 5.2).

These timelines illustrate how the technology base is "pulled" by functional requirements passed to it by the military applications (See App. I). They also show how the technology base is open to exploitation of discoveries and technologies produced under ongoing 6.1 (basic research) programs.

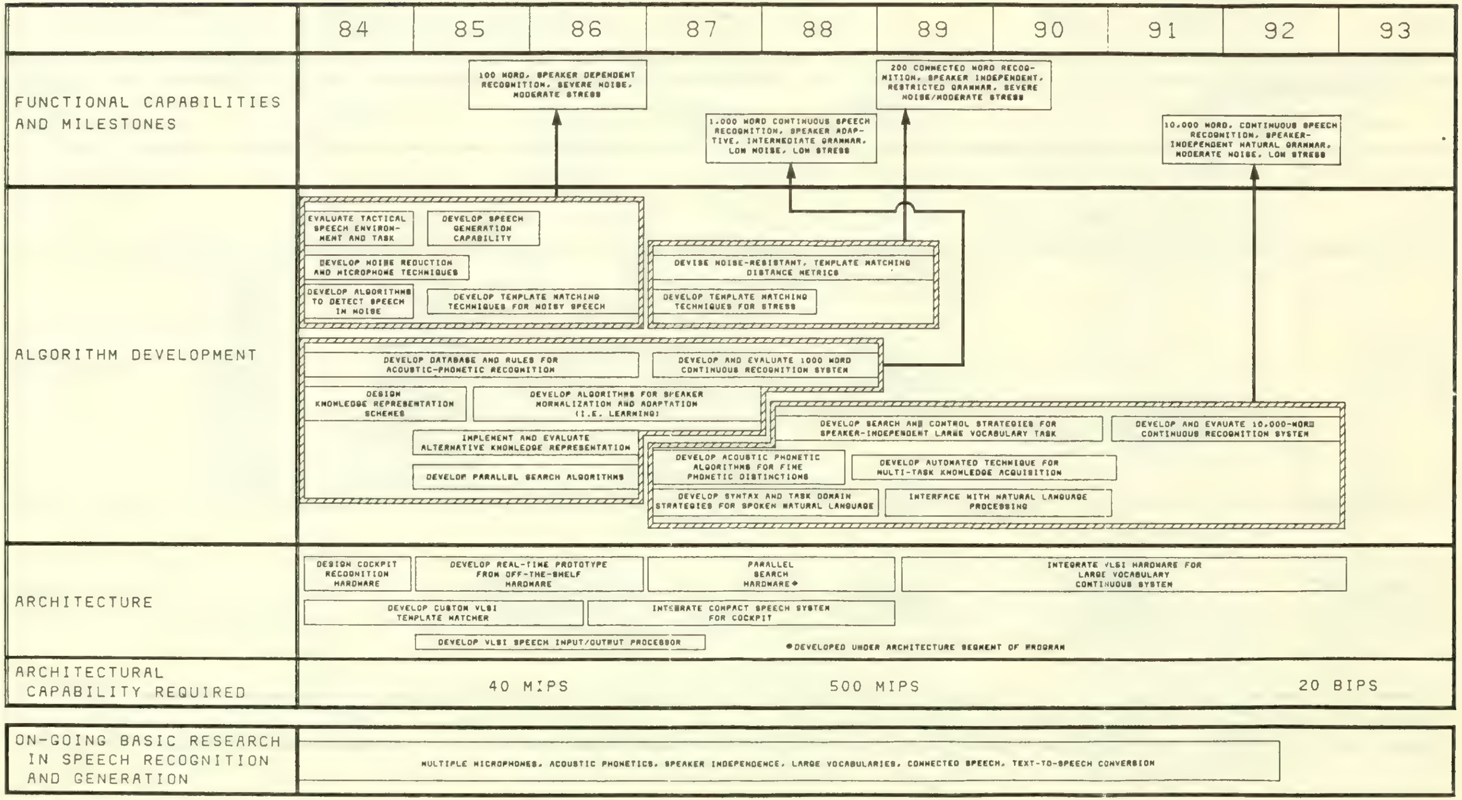
VISION SUBSYSTEMS

VERSION OF 10-18-83 II.1.1



SPEECH SUBSYSTEMS

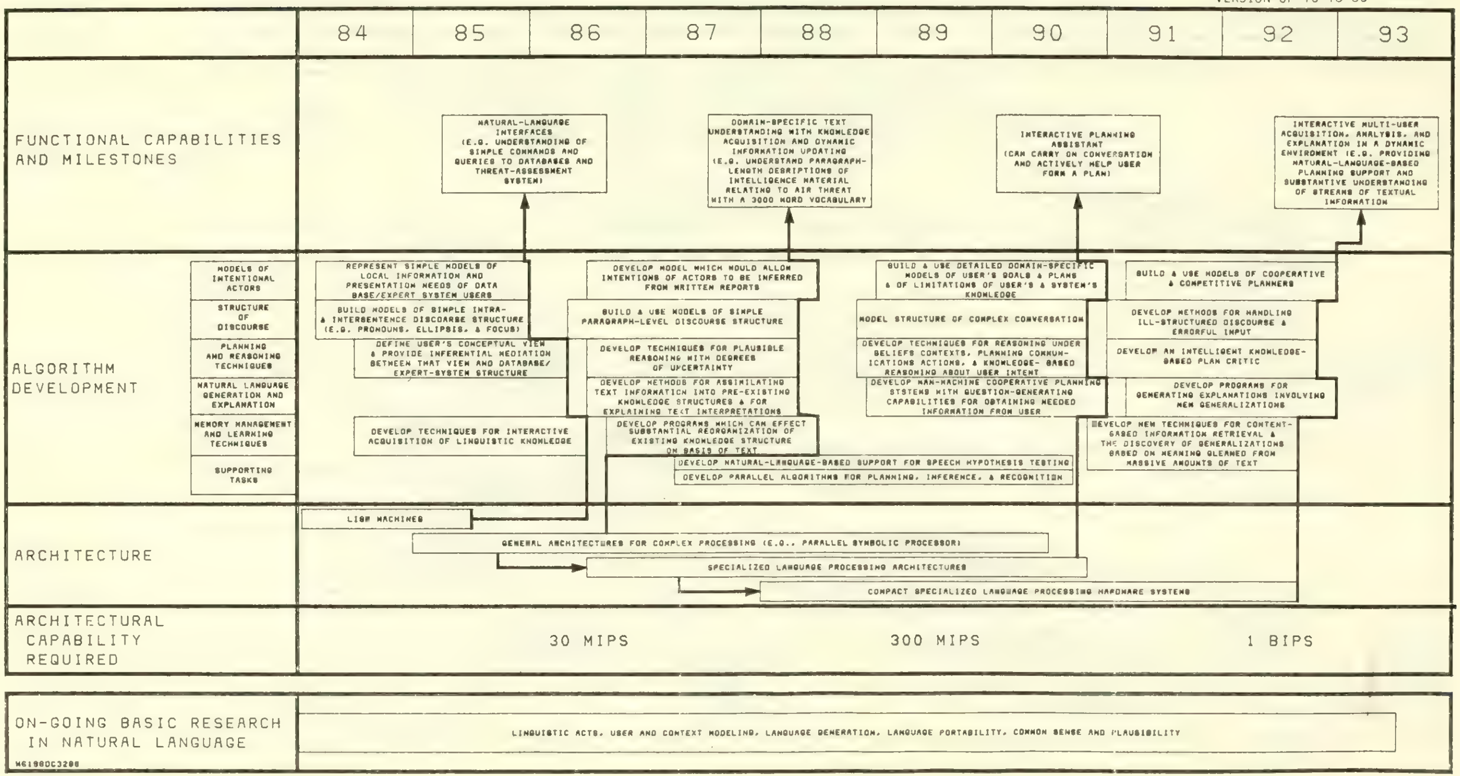
VERSION OF 10-18-83 II.1.2

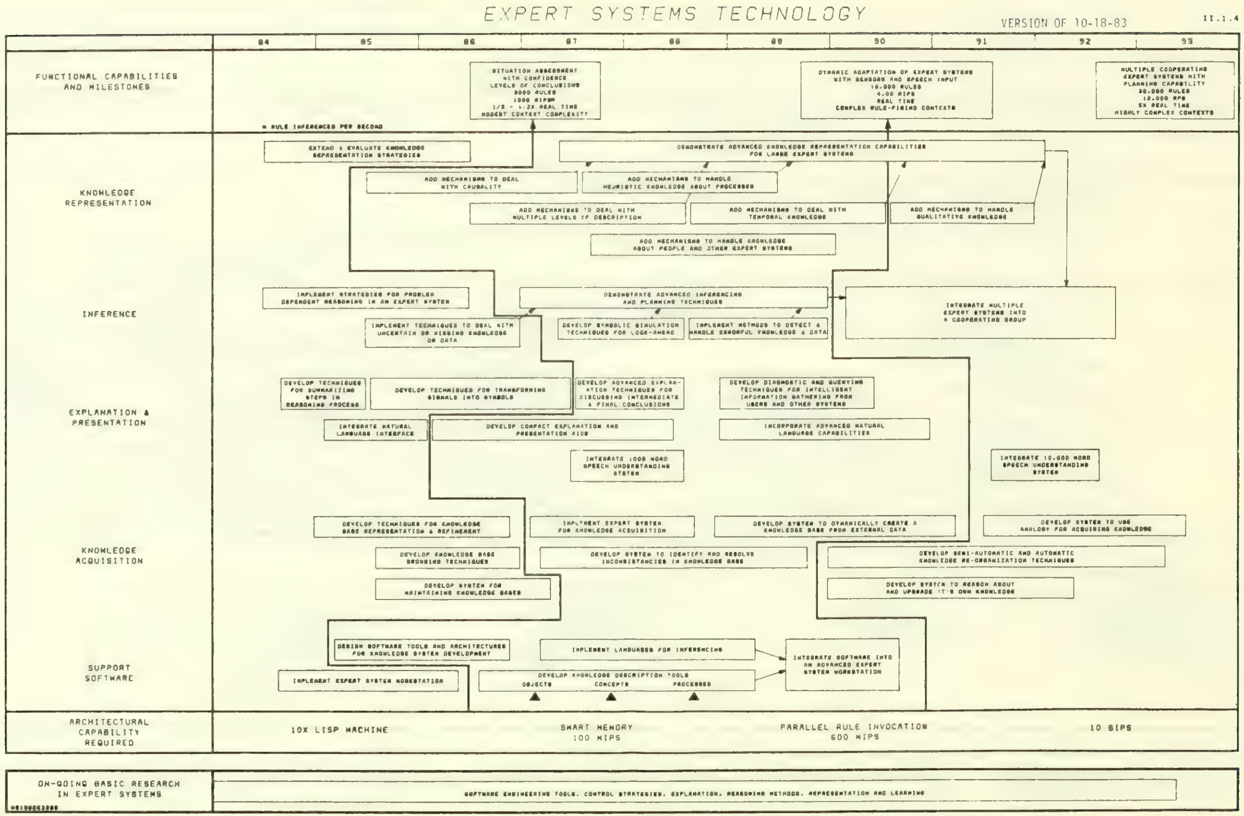


M61988E3286

NATURAL LANGUAGE SUBSYSTEMS

VERSION OF 10-18-83 II.1.3

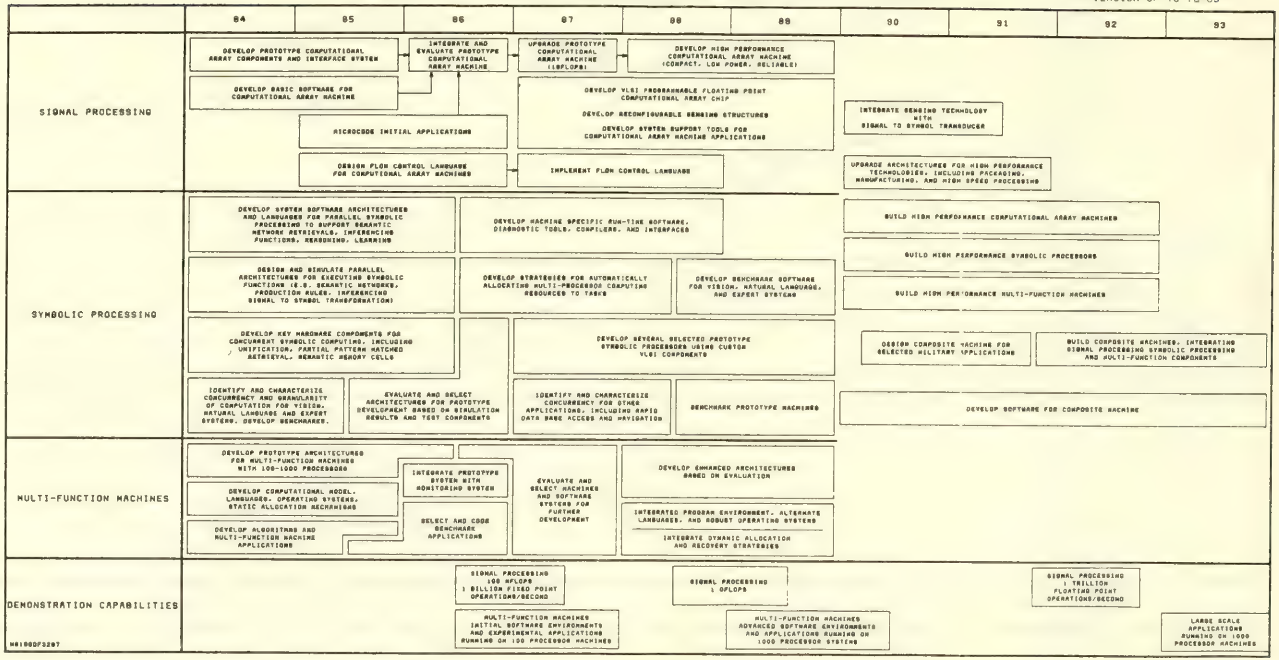



EXPERT SYSTEMS TECHNOLOGY
VERSION OF 10-18-83
II.1.4
84
85
86
87
88
89
90
91
92
93
FUNCTIONAL CAPABILITIES AND MILESTONES
SITUATION ASSESSMENT WITH CONFIDENCE LEVELS OF CONCLUSIONS 3000 RULES 1000 RIPS* 1/3 - 1/2X REAL TIME MODEST CONTEXT COMPLEXITY
DYNAMIC ADAPTATION OF EXPERT SYSTEMS WITH SENSORS AND SPEECH INPUT 10,000 RULES 4,00 RIPS REAL TIME COMPLEX RULE-FIRING CONTEXTS
MULTIPLE COOPERATING EXPERT SYSTEMS WITH PLANNING CAPABILITY 30,000 RULES 12,000 RPS 5X REAL TIME HIGHLY COMPLEX CONTEXTS
* RULE INFERENCES PER SECOND
KNOWLEDGE REPRESENTATION
EXTEND & EVALUATE KNOWLEDGE REPRESENTATION STRATEGIES
DEMONSTRATE ADVANCED KNOWLEDGE REPRESENTATION CAPABILITIES FOR LARGE EXPERT SYSTEMS
ADD MECHANISMS TO DEAL WITH CAUSALITY
ADD MECHANISMS TO HANDLE HEURISTIC KNOWLEDGE ABOUT PROCESSES
ADD MECHANISMS TO DEAL WITH MULTIPLE LEVELS OF DESCRIPTION
ADD MECHANISMS TO DEAL WITH TEMPORAL KNOWLEDGE
ADD MECHANISMS TO HANDLE QUALITATIVE KNOWLEDGE
ADD MECHANISMS TO HANDLE KNOWLEDGE ABOUT PEOPLE AND OTHER EXPERT SYSTEMS
INFERENCE
IMPLEMENT STRATEGIES FOR PROBLEM DEPENDENT REASONING IN AN EXPERT SYSTEM
DEMONSTRATE ADVANCED INFERENCING AND PLANNING TECHNIQUES
INTEGRATE MULTIPLE EXPERT SYSTEMS INTO A COOPERATING GROUP
IMPLEMENT TECHNIQUES TO DEAL WITH UNCERTAIN OR MISSING KNOWLEDGE OR DATA
DEVELOP SYMBOLIC SIMULATION TECHNIQUES FOR LOOK-AHEAD
IMPLEMENT METHODS TO DETECT & HANDLE ERRORFUL KNOWLEDGE & DATA
EXPLANATION & PRESENTATION
DEVELOP TECHNIQUES FOR SUMMARIZING STEPS IN REASONING PROCESS
DEVELOP TECHNIQUES FOR TRANSFORMING SIGNALS INTO SYMBOLS
DEVELOP ADVANCED EXPLANATION TECHNIQUES FOR DISCUSSING INTERMEDIATE & FINAL CONCLUSIONS
DEVELOP DIAGNOSTIC AND QUERYING TECHNIQUES FOR INTELLIGENT INFORMATION GATHERING FROM USERS AND OTHER SYSTEMS
INTEGRATE NATURAL LANGUAGE INTERFACE
DEVELOP COMPACT EXPLANATION AND PRESENTATION AIDS
INCORPORATE ADVANCED NATURAL LANGUAGE CAPABILITIES
INTEGRATE 1000 WORD SPEECH UNDERSTANDING SYSTEM
INTEGRATE 10,000 WORD SPEECH UNDERSTANDING SYSTEM
KNOWLEDGE ACQUISITION
DEVELOP TECHNIQUES FOR KNOWLEDGE BASE REPRESENTATION & REFINEMENT
IMPLEMENT EXPERT SYSTEM FOR KNOWLEDGE ACQUISITION
DEVELOP SYSTEM TO DYNAMICALLY CREATE A KNOWLEDGE BASE FROM EXTERNAL DATA
DEVELOP SYSTEM TO USE ANALOGY FOR ACQUIRING KNOWLEDGE
DEVELOP KNOWLEDGE BASE BROWSING TECHNIQUES
DEVELOP SYSTEM TO IDENTIFY AND RESOLVE INCONSISTENCIES IN KNOWLEDGE BASE
DEVELOP SEMI-AUTOMATIC AND AUTOMATIC KNOWLEDGE RE-ORGANIZATION TECHNIQUES
DEVELOP SYSTEM FOR MAINTAINING KNOWLEDGE BASES
DEVELOP SYSTEM TO REASON ABOUT AND UPGRADE IT'S OWN KNOWLEDGE
SUPPORT SOFTWARE
DESIGN SOFTWARE TOOLS AND ARCHITECTURES FOR KNOWLEDGE SYSTEM DEVELOPMENT
IMPLEMENT LANGUAGES FOR INFERENCING
INTEGRATE SOFTWARE INTO AN ADVANCED EXPERT SYSTEM WORKSTATION
IMPLEMENT EXPERT SYSTEM WORKSTATION
DEVELOP KNOWLEDGE DESCRIPTION TOOLS
OBJECTS
CONCEPTS
PROCESSES
ARCHITECTURAL CAPABILITY REQUIRED
10X LISP MACHINE
SMART MEMORY 100 MIPS
PARALLEL RULE INVOCATION 600 MIPS
10 BIPS
ON-GOING BASIC RESEARCH IN EXPERT SYSTEMS
SOFTWARE ENGINEERING TOOLS, CONTROL STRATEGIES, EXPLANATION, REASONING METHODS, REPRESENTATION AND LEARNING

MACHINE HARDWARE/SOFTWARE ARCHITECTURE

VERSION OF 10-18-83

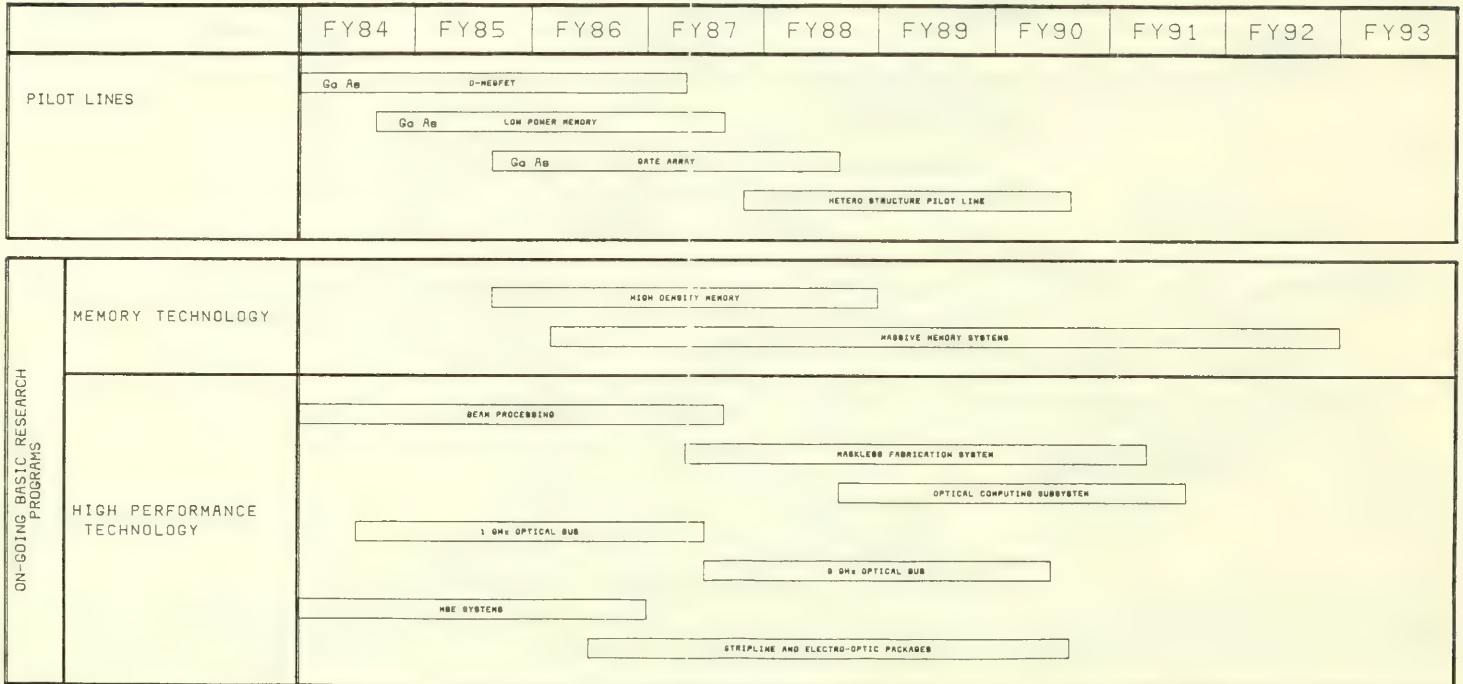
II.2



VERSION OF 10-18-83

SUPPORTING MICROELECTRONICS TECHNOLOGY

II.3



W6196BM3277

APPENDIX III

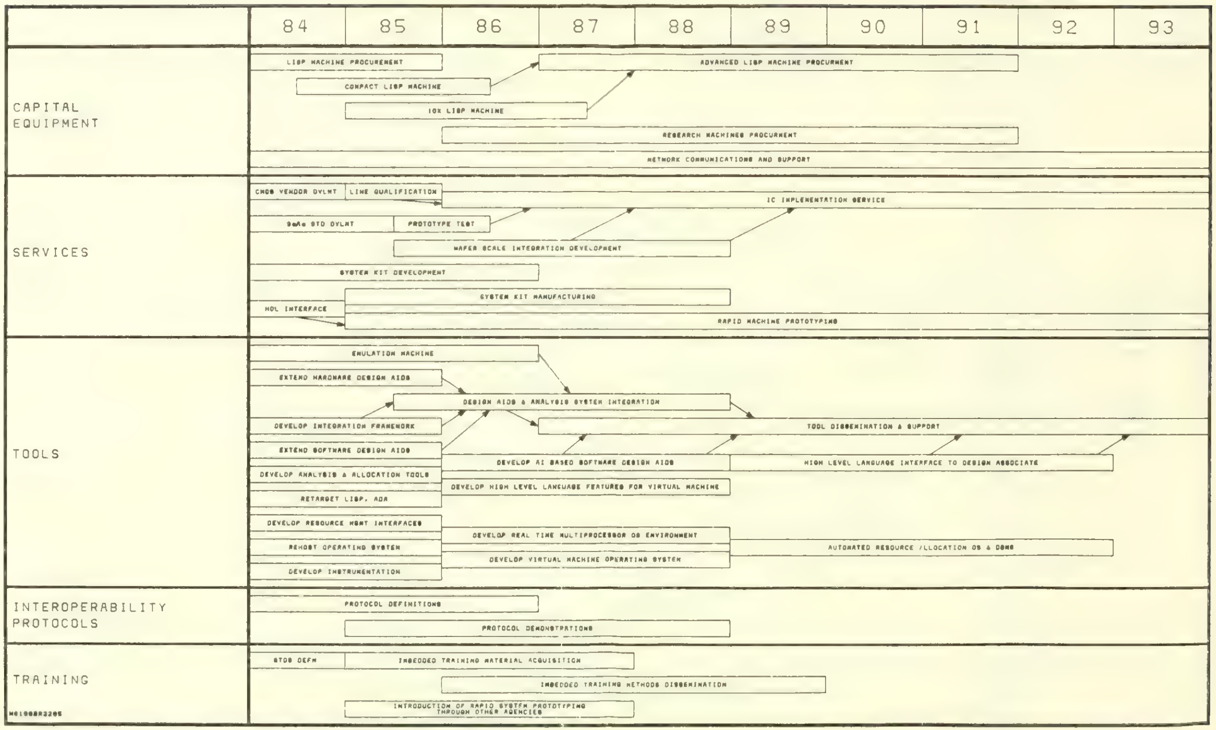
INFRASTRUCTURE PLANS

This appendix contains plans and timelines for the Strategic Computing program infrastructure (see also Chapter 5, Section 5.3).

INFRASTRUCTURE

III.0

VERSION OF 10-18-83



APPENDIX IV

PROGRAM WORK BREAKDOWN STRUCTURE

This appendix contains a detailed breakdown of the program work structure used for planning and budgeting.

C.1

1.0 MILITARY APPLICATIONS

1.1 AUTONOMOUS VEHICLES
- 1.1.1 LAND VEHICLE
  - 1.1.1.1 REQUIREMENTS ANALYSIS
  - 1.1.1.2 DEMONSTRATION SYSTEM DESIGN
  - 1.1.1.3 SYSTEM INTEGRATION
  - 1.1.1.4 FUNCTIONAL TEST
  - 1.1.1.5 FURNISHED EQUIPMENT
- 1.1.2 SUBMARINE VEHICLE
  - 1.1.2.1 REQUIREMENTS ANALYSIS
  - 1.1.2.2 DEMONSTRATION SYSTEM DESIGN
  - 1.1.2.3 SYSTEM INTEGRATION
  - 1.1.2.4 FUNCTIONAL TEST
  - 1.1.2.5 FURNISHED EQUIPMENT
- 1.1.3 AIR VEHICLE
  - 1.1.3.1 REQUIREMENTS ANALYSIS
  - 1.1.3.2 DEMONSTRATION SYSTEM DESIGN
  - 1.1.3.3 SYSTEM INTEGRATION
  - 1.1.3.4 FUNCTIONAL TEST
  - 1.1.3.5 FURNISHED EQUIPMENT
- 1.1.4 SPACE VEHICLE
  - 1.1.4.1 REQUIREMENTS ANALYSIS
  - 1.1.4.2 DEMONSTRATION SYSTEM DESIGN
  - 1.1.4.3 SYSTEM INTEGRATION
  - 1.1.4.4 FUNCTIONAL TEST
  - 1.1.4.5 FURNISHED EQUIPMENT

1.2 OPERATIONAL ASSOCIATES
- 1.2.1 PILOT'S ASSOCIATE
  - 1.2.1.1 REQUIREMENTS ANALYSIS
  - 1.2.1.2 DEMONSTRATION SYSTEM DESIGN
  - 1.2.1.3 SYSTEM INTEGRATION
  - 1.2.1.4 FUNCTIONAL TEST
  - 1.2.1.5 FURNISHED EQUIPMENT
- 1.2.2 TANK CREW'S ASSOCIATE
  - 1.2.2.1 REQUIREMENTS ANALYSIS
  - 1.2.2.2 DEMONSTRATION SYSTEM DESIGN
  - 1.2.2.3 SYSTEM INTEGRATION
  - 1.2.2.4 FUNCTIONAL TEST
  - 1.2.2.5 FURNISHED EQUIPMENT

1.3 BATTLE MANAGEMENT
- 1.3.1 BATTALION BATTLE MANAGEMENT
  - 1.3.1.1 REQUIREMENTS ANALYSIS
  - 1.3.1.2 DEMONSTRATION SYSTEM DESIGN
  - 1.3.1.3 SYSTEM INTEGRATION
  - 1.3.1.4 FUNCTIONAL TEST
  - 1.3.1.5 FURNISHED EQUIPMENT
- 1.3.2 FLEET BATTLE MANAGEMENT
  - 1.3.2.1 REQUIREMENTS ANALYSIS
  - 1.3.2.2 DEMONSTRATION SYSTEM DESIGN
  - 1.3.2.3 SYSTEM INTEGRATION
  - 1.3.2.4 FUNCTIONAL TEST
  - 1.3.2.5 FURNISHED EQUIPMENT
- 1.3.3 BALLISTIC MISSILE DEFENSE
  - 1.3.3.1 REQUIREMENTS ANALYSIS
  - 1.3.3.2 DEMONSTRATION SYSTEM DESIGN
  - 1.3.3.3 SYSTEM INTEGRATION
  - 1.3.3.4 FUNCTIONAL TEST
  - 1.3.3.5 FURNISHED EQUIPMENT
- 1.3.4 ADAPTIVE ELECTRONIC WARFARE
  - 1.3.4.1 REQUIREMENTS ANALYSIS
  - 1.3.4.2 DEMONSTRATION SYSTEM DESIGN
  - 1.3.4.3 SYSTEM INTEGRATION
  - 1.3.4.4 FUNCTIONAL TEST
  - 1.3.4.5 FURNISHED EQUIPMENT

W6196CA3278

WORK BREAKDOWN:
MILITARY APPLICATIONS

C.2

2.0 TECHNOLOGY BASE

2.1 INTEGRATED INTELLIGENT FUNCTIONS
- 2.1.1 VISION
  - 2.1.1.1 GENERAL
  - 2.1.1.2 MISSION SPECIFIC
    - 2.1.1.2.1 AUTONOMOUS VEHICLE
- 2.1.2 SPEECH
  - 2.1.2.1 GENERAL
  - 2.1.2.2 MISSION SPECIFIC
    - 2.1.2.2.1 PILOT'S ASSOCIATE
- 2.1.3 NATURAL LANGUAGE
  - 2.1.3.1 GENERAL
  - 2.1.3.2 MISSION SPECIFIC
    - 2.1.3.2.1 BATTLE MANAGEMENT
- 2.1.4 EXPERT SYSTEMS
  - 2.1.4.1 GENERAL
  - 2.1.4.2 MISSION SPECIFIC
    - 2.1.4.2.1 AUTONOMOUS VEHICLE
    - 2.1.4.2.2 PILOT'S ASSOCIATE
    - 2.1.4.2.3 BATTLE MANAGEMENT
- 2.1.5 APPLICATION SPECIFIC
  - 2.1.5.1 SIGNAL INTERPRETATION
  - 2.1.5.2 INFORMATION FUSION/MACHINE LEARNING
  - 2.1.5.3 PLANNING AND REASONING
  - 2.1.5.4 KNOWLEDGE AND DATA MANAGEMENT
  - 2.1.5.5 SIMULATION, MODELING, AND CONTROL
  - 2.1.5.6 NAVIGATION
  - 2.1.5.7 GRAPHICS DISPLAY/IMAGE GENERATION
  - 2.1.5.8 DISTRIBUTED COMMUNICATIONS

2.2 SYSTEM ARCHITECTURES
- 2.2.1 SIGNAL PROCESSORS
  - 2.2.1.1 GENERAL
    - 2.2.1.1.1 COMPUTATIONAL ARRAYS
    - 2.2.1.1.2 RECONFIGURABLE SENSING STRUCTURES
  - 2.2.1.2 MISSION SPECIFIC
- 2.2.2 SYMBOLIC PROCESSORS
  - 2.2.2.1 GENERAL
    - 2.2.2.1.1 SIGNAL TO SYMBOL TRANSDUCER
    - 2.2.2.1.2 SEMANTIC MEMORY ENGINE
    - 2.2.2.1.3 PRODUCTION RULE MACHINE
    - 2.2.2.1.4 FUSION MACHINE
    - 2.2.2.1.5 INFERENCE & CONTROL MACHINE
    - 2.2.2.1.6 SEARCH MACHINE
  - 2.2.2.2 MISSION SPECIFIC
- 2.2.3 GENERAL PURPOSE
  - 2.2.3.1 GENERAL
    - 2.2.3.1.1 MULTI-MICROPROCESSOR
    - 2.2.3.1.2 DATA FLOW MACHINE
  - 2.2.3.2 MISSION SPECIFICS
- 2.2.4 FUNCTION SPECIFIC
  - 2.2.4.1 DATA BASE MACHINE
  - 2.2.4.2 SIMULATION MACHINE
  - 2.2.4.3 DISPLAY MACHINE

2.3 MICROELECTRONICS
- 2.3.1 GAAS D-MESFET PILOT LINE
- 2.3.2 GAAS LOW POWER MEMORY PILOT LINE
- 2.3.3 GAAS GATE ARRAY PILOT LINE
- 2.3.4 HETEROSTRUCTURES PILOT LINE

H6916C83286

WORK BREAKDOWN:
TECHNOLOGY BASE

3.0 DEVELOPMENT INFRASTRUCTURE

- 3.1 FACILITIES
  - 3.1.1 LISP MACHINES
  - 3.1.2 RESEARCH MACHINES
  - 3.1.3 COMMUNICATION NETWORKS
- 3.2 SERVICES
  - 3.2.1 INTEGRATED CIRCUIT IMPLEMENTATION
  - 3.2.2 RAPID MACHINE PROTOTYPING
  - 3.2.3 SYSTEM INTEROPERABILITY KITS
- 3.3 TOOLS
  - 3.3.1 SYSTEM TOOLS
  - 3.3.2 EMULATION FACILITIES
  - 3.3.3 MULTIPROCESSOR SOFTWARE WORKBENCH
  - 3.3.4 PROGRAMMING ENVIRONMENT
- 3.4 STANDARDS
  - 3.4.1 INTEROPERABILITY PROTOCOLS
- 3.5 OTHER
  - 3.5.1 TRAINING AND WORKSHOPS
  - 3.5.2 LABORATORY AIDS

W6196CC3277

WORK BREAKDOWN:
DEVELOPMENT INFRASTRUCTURE

C.4

4.0 PROGRAM OFFICE SUPPORT

- 4.1 PROGRAM MANAGEMENT
- 4.2 TECHNICAL SUPPORT
  - 4.2.1 BENCHMARKING
  - 4.2.2 TECHNICAL LIAISON
  - 4.2.3 TECHNICAL EVALUATION FOR SOURCE SELECTION
- 4.3 INFORMATION MANAGEMENT
  - 4.3.1 DOCUMENTATION
  - 4.3.2 MIS AND MANAGEMENT TOOLS

W6196CD3277

WORK BREAKDOWN:
PROGRAM OFFICE SUPPORT

APPENDIX V
COLLATERAL ACTIVITIES

Within the United States and abroad, there are current and planned technology programs that relate directly to the goals and activities of the Strategic Computing Program. They have been reviewed and a list and description of the activities compiled. Many relate to proprietary commercial activities and to classified military programs. Access to this data will be available to those with proper clearances and need to know.

STRATEGIC COMPUTING

First Annual Report

New-Generation Computing Technology:
A Strategic Plan for its Development
and Application to Critical
Problems in Defense

February 1985

DEFENSE ADVANCED RESEARCH PROJECTS AGENCY STRATEGIC COMPUTING PROGRAM FIRST ANNUAL REPORT

DIRECTOR'S COMMENTS:

"In the first year of operation, the Strategic Computing Program, which was started in November of 1983 at the Defense Advanced Research Projects Agency (DARPA), has made major progress toward program goals. Those who have been associated with the planning and initial phase of this initiative rate it among the most important research programs ever undertaken at DARPA.

There are four fundamental goals in the overall program. First, and foremost, is to advance machine intelligence technology across a broad front to maintain with assurance the U.S. technical lead in advanced computer technology through the next decade. To accomplish this goal we have placed emphasis on the Artificial Intelligence disciplines of 1) speech recognition and understanding, 2) natural language computer interfaces, 3) vision comprehension systems, and 4) advanced expert systems development. To provide the 100 to 10,000 fold increases in computer performance required to support meaningful military applications of machine intelligence, many advanced multiprocessor computer architectures and advanced microelectronics techniques are being pursued. Several military applications testbeds are being set up with significant Army, Navy and Air Force problems emphasized.

The second major goal is to transition this technology from prior and currently DARPA-sponsored university research efforts to the military services and the Defense industry which serves them. Close cooperation with the military departments and competitive research contracts, with industry and universities jointly participating, are the mechanisms being used.

A third program goal is to increase the availability of AI-trained scientists and engineers in our universities through substantially increased funding of sponsored graduate student research in Artificial Intelligence. Between FY84 and FY85 alone, a 100% increase in student participation is anticipated.

The final goal is to provide a broad base of supporting research for advanced machine intelligence technology through adequate elements of infrastructure such as advanced networking, rapid turnaround microcircuit fabrication facilities, advanced list processors, large-scale emulation facilities and broad research worker access to new computing machines as they become available.

Progress in all areas to date has been gratifying, and the program remains in a period of rapid acceleration with high expectations of success."

Robert S. Cooper
Director

ACCOMPLISHMENTS

AUTONOMOUS LAND VEHICLE TESTBED

AUTONOMOUS LAND VEHICLE

The Autonomous Land Vehicle (ALV) is a technology demonstration vehicle capable of independent operation in complex terrain. Upon receiving tasking information, the vehicle, through vision systems, image understanding, and expert route planning systems, will execute a navigation plan to accomplish its mission. Martin-Marietta of Denver has been selected, on a competitive basis, as the ALV integration contractor. In this role, Martin-Marietta will integrate the work of several university and industrial contractors. Martin-Marietta has designed the vehicle testbed environment and has formulated the architecture of the software modules for the vehicle and at the Remote Computing Site, where actual vehicle operations will be conducted. The vehicle will serve as a testbed environment accessible to researchers across the country for the evaluation of image understanding algorithms under actual vehicle operating conditions. Full motion color video-teleconferencing is being employed to link key research organizations with the vehicle testbed environment to provide near-realtime observation and analysis of the experiments being performed.

NAVAL BATTLE MANAGEMENT

NAVAL BATTLE MANAGEMENT

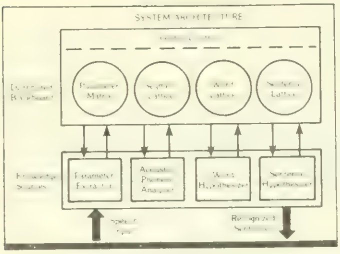
The Naval Battle Management program consists of enhanced command assistance at the Fleet Command Center and Carrier Battle Group levels. The Fleet Command Center Battle Management Program (FCCBMP) will develop an advanced computer architecture and expert system testbed at the Pacific Fleet Command Center. Texas Instruments will develop the Force Requirements Expert System (FRESH), the first of five expert systems to be developed by the FCCBMP. FRESH will assist in the monitoring of force readiness and will recommend forces to respond to fleet contingencies. A preliminary situation assessment expert system for evaluation on-board the aircraft carrier U.S.S. Carl Vinson (Carrier Battle Group application), has been developed by Carnegie-Mellon University, with the Computer Corporation of America developing a graphic interface to the system for enhanced user-system interaction.

ROAD FOLLOWING

VISION BASED NAVIGATION

The University of Maryland is developing vision algorithms for the road following and obstacle avoidance demonstrations, using the ALV testbed. The vision techniques being pursued are a combination of image segmentation and knowledge-based extraction of features. Initial algorithm milestones have been met, and the University of Maryland has developed a robotic arm simulation to test and evaluate the vision software. A representative of Martin-Marietta, the ALV integrating contractor, is in residence at Maryland to facilitate technology transition to the ALV testbed.

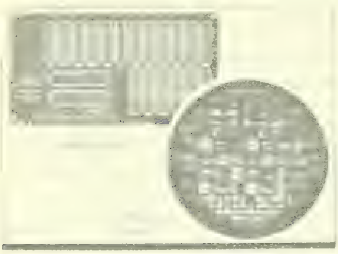
ACCOMPLISHMENTS (Continued)



SPEECH RECOGNITION

SPEECH RECOGNITION

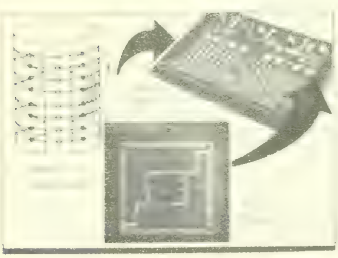
Carnegie-Mellon University is developing a system that will enable a computer to understand continuous speech as it is spoken by different people with varying dialects and speech patterns. Initially the system will utilize a 1,000 word vocabulary, with the goal of a 10,000 word vocabulary when fully developed. A multiprocessor system will be used to provide real-time speech recognition performance. The system architecture will permit the integration of multiple knowledge sources that include speech wave-form analysis, identification of specific features of the speech utterance using acoustic-phonetic techniques, and incorporation of natural language processing capabilities to extend power and flexibility. A functional simulation of the system has been constructed and tested in order to focus the research on issues of modular knowledge-source integration and distributed processing implementation.



RAPIDLY PROTOTYPED INTEGRATED CIRCUITS AND PRINTED CIRCUIT BOARD

MOS IMPLEMENTATION SERVICE

The MOS Implementation Service (MOSIS) is a vehicle whereby researchers and designers across the country can rapidly acquire prototype custom-designed integrated circuits and printed circuit boards. MOSIS provides a mechanism for access to state-of-the-art fabrication lines at low cost through cost sharing using multiproject wafers. Circuit specifications are transmitted via the ARPANET communications network to the MOSIS Center at USC-ISI, which coordinates manufacture, test, and product delivery. Manufacturing times from receipt of data at MOSIS to shipment of assembled parts in as little as four weeks have been obtained. The availability of the MOSIS service to the research community provides a unique capability to implement new ideas in VLSI design techniques as well as novel system architectures.

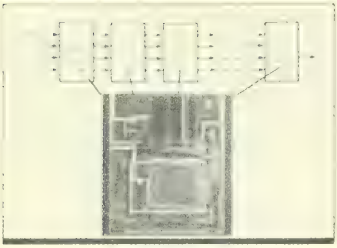


BUTTERFLY SWITCH NETWORK

BUTTERFLY MULTIPROCESSOR

The Bolt, Beranek, and Newman (BBN) Inc. Butterfly Multiprocessor is the first large-scale multiprocessor architecture to emerge from the Strategic Computing program. Based upon DARPA-sponsored work in developing a switching technology for wide-band communication, the Butterfly architecture is implemented using commercial microprocessors for the processing elements and a custom-designed VLSI switch circuit fabricated via MOSIS. A 128 node machine will be developed in early FY85 and be accessible via the ARPANET communications network. A number of smaller research machines will be delivered late in FY85 for multiprocessor software development and evaluation in the applications environment.

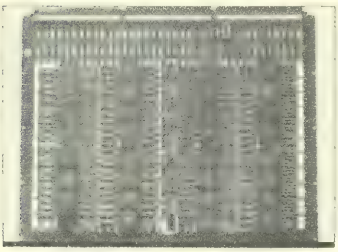
ACCOMPLISHMENTS (Continued)



WARP MACHINE

PROGRAMMABLE SYSTOLIC ARRAY

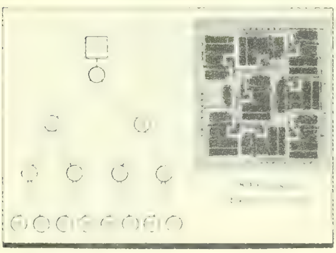
A systolic array is a configuration of multiple processing elements which rapidly compute complex mathematical operations such as the fast Fourier transform and convolution operators. A prototype system has been built which incorporates nine experimental microprogrammable integrated circuits designed by Carnegie Mellon University, and prototyped through MOSIS. This prototype operates at a rate of 10 million instructions per second (MIPS). The next phase of the program will use commercially available microprocessors interconnected by a custom-designed VLSI circuit. The second prototype, the WARP machine, will incorporate 10 cells, each operating at a rate of 10 million floating point operations per second (MFLOPS) to provide a 100 MFLOPS capability. The WARP will be demonstrated in late 1985



THE CONNECTION MACHINE
512 NODE PROCESSOR BOARD

CONNECTION MACHINE

The Connection Machine is a very fine-grain parallel computer architecture designed for machine intelligence applications requiring processing capabilities that are several orders of magnitude greater than those of current state-of-the-art computers. The 64 thousand processor prototype connection machine will have approximately 1000 times the logical inference performance capabilities of current LISP (List Processor) workstations. Processors are connected so that each processor may communicate with any other through a fast message routing system. The initial 64 thousand node processor machines will be fabricated using conservative VLSI technology of 10,000 gate CMOS gate arrays. Fully functional processor chips, each containing 16 processor elements and a communications router, have been demonstrated. Future Connection Machines will be designed with a custom VLSI circuit suitable for use in a 1 million processor machine. The Connection Machine prototypes will be accessed as shared memory from a host computer together with an instruction stream generator. Operating in a LISP environment, the Connection Machine will be accessible to the research community via the ARPANET.

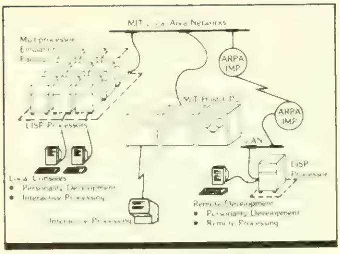


TREE MACHINE ARRAY

TREE MACHINES

Columbia University has designed and is building two tree structured multiprocessor architectures called the DADO and the NON-VON. The NON-VON is a fine-grain, binary tree architecture. Four versions of the NON-VON have been designed and one small-scale prototype built. The DADO is a coarse-grain tree architecture which has been designed for parallel processing in a production rule environment. The project includes design, test, and evaluation of custom VLSI circuits and printed circuit boards fabricated through MOSIS, as well as development of supporting software and implementation of higher level programming languages.

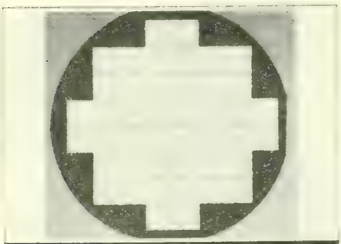
ACCOMPLISHMENTS (Continued)



DATAFLOW EMULATION FACILITY

DATAFLOW EMULATION FACILITY

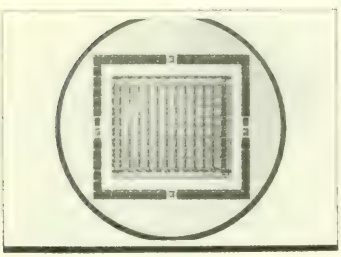
A Multiprocessor Emulation Facility for tagged token dataflow architectures has been designed and is being developed at the Massachusetts Institute of Technology Laboratory for Computer Science. The facility will use up to 64 commercially available LISP machines connected by a high-bandwidth low latency switching network. A small-scale hardware and software system has been completed. The design of the high performance interconnected system has also been completed.



GaAs VLSI WAFER

GaAs PILOT LINE

The successful implementation of the next generation of advanced computer technologies is critically dependent upon parallel successes in both hardware and software. To meet the emerging requirements for high speed and very low power VLSI circuits demanded by the information processing rates of future systems, DARPA initiated the development of a Gallium Arsenide (GaAs) Pilot Line as part of the 1984 Strategic Computing Program. Rockwell has been selected to establish a pilot line capable of processing 100 wafer starts per week by 1987, using a 6,000 gate array and a 16K RAM as the demonstration vehicles. Honeywell, under a subcontract to Rockwell, will establish a compatible pilot line facility. With the enhanced radiation hardness demonstrated for the GaAs technology, the responsibility for the Pilot Line Program was transferred to the Strategic Defense Initiative at the beginning of FY85. In concert with the Infrastructure program of Strategic Computing, access to the GaAs Pilot Line will become available to the research community via the MOSIS Service.



LASER PROGRAMMED WAFER

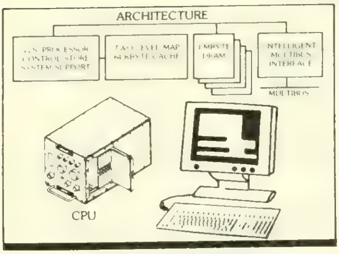
WAFER SCALE INTEGRATION

The Adaptive Wafer Scale Integration program at MIT/Lincoln Laboratories provides a rapid prototyping capability for highly complex integrated circuits. By using a computer-controlled laser to connect or open fusible links on the wafer it is possible to rapidly configure custom circuits by interconnecting a series of common circuit building blocks. The laser-link technology is now being integrated into the MOSIS system. Also being developed is a VLSI compatible electron-beam programmable interconnect technology which will allow individual wafers to be reconfigured to incorporate design changes with 24-hour turnaround.

ACCOMPLISHMENTS (Continued)

COMPACT LISP MACHINE

Texas Instruments (TI) is developing a Compact LISP machine based upon high-performance semiconductor technology. A VLSI LISP processor chip, fabricated in 2 micron CMOS technology, has been designed for the central processing unit (CPU). In addition, the computer will incorporate advanced high-speed static RAM memory chips, developed under TI's VHSIC program. The high-complexity VLSI chip is designed to operate at speeds up to 40 MHz, creating a LISP processor that performs, on one chip, functions requiring several hundred integrated circuits in current computer systems. The new chip is designed to provide two to ten times the processing power of today's commercial symbolic processors.

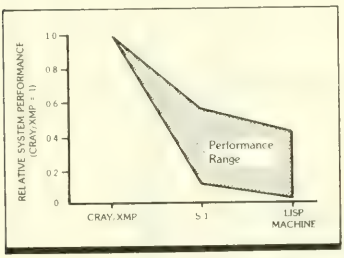


COMPACT LISP MACHINE

BENCHMARKING

DARPA has sponsored a project at Stanford University to characterize implementations of the LISP language on existing uni- and multiprocessor systems. Analyses of specific algorithms addressed the impact of basic hardware configuration, LISP "instructions," LISP "functions," and major LISP system facilities on overall system performance. Efficient use of a given machine is also dependent upon programming flexibility and software development tools.

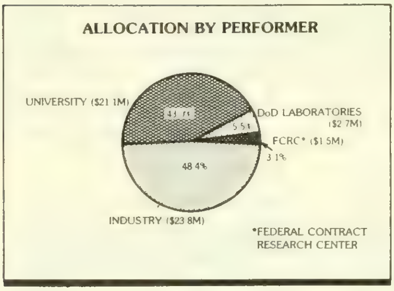
Results were presented at the 1984 meeting of the American Association for Artificial Intelligence. One interesting result of the project was that the Cray-XMP supercomputer, a machine designed for numerical calculations, performed symbolic operations exceptionally well. The fundamental understanding of the benchmarking methodology that has resulted from this project forms the basis for the characterization and analysis of the multiprocessor system architectures that will emerge from the Strategic Computing Program.

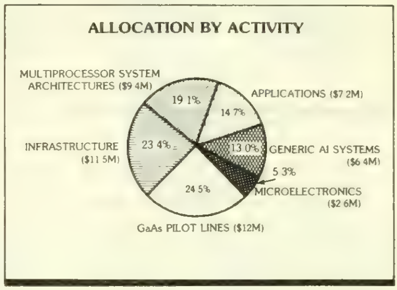


COMPARISON OF SYSTEM BENCHMARKS

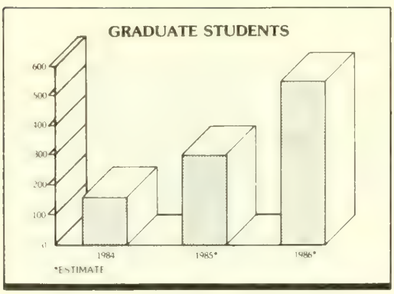
STRATEGIC COMPUTING BUDGET ALLOCATION
— Fiscal Year 1984 —

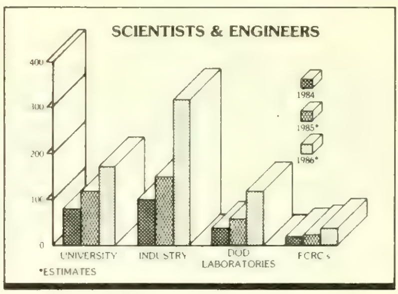
Total Budget: $49.1 Million





PERSONNEL SUPPORTED BY STRATEGIC COMPUTING

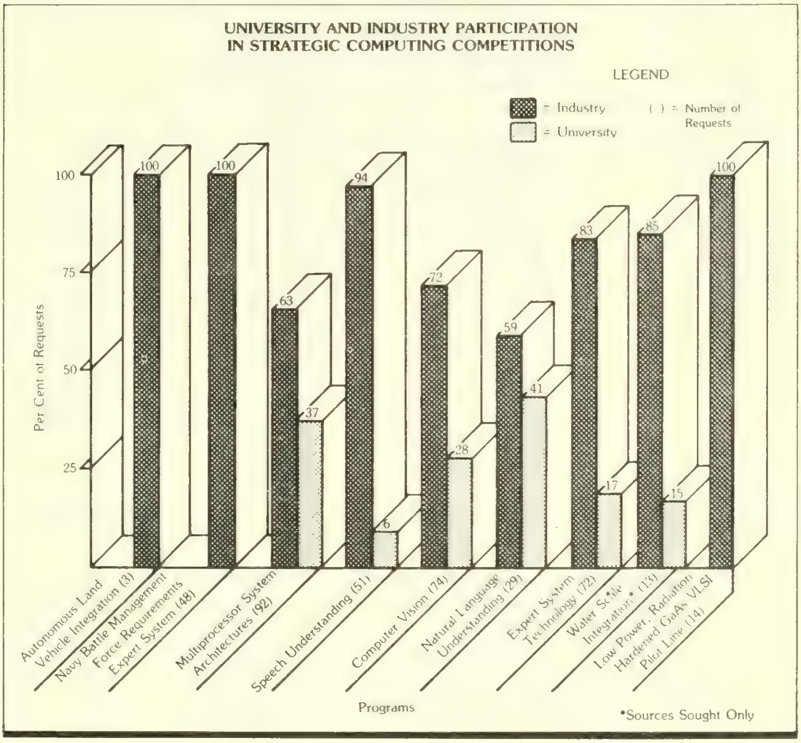




COMPETITIVE PROCUREMENT

Several procedures for competitive procurement have been used within the first year of the Strategic Computing Program, consistent with Public Law 98-72. In particular, about half of the funds have been allocated through the issuance of Requests for Proposals (RFPs). The result has been to enable DARPA to involve the highest quality university and industry researchers in the Program; to promote a transition of leading edge ideas from DARPA-supported research efforts to major defense contractors, primarily through joint industry-university proposals; and to increase the diversity of performers within the Program by including organizations that have not traditionally worked with DARPA.

The key competitive procurements initiated in the past year are shown in the figure below. The relative participation by the university and industry communities is shown for these procurements.



PROJECT AWARDS

ACTIVITY	PROJECT	PERFORMER	AGENT
AUTONOMOUS LAND VEHICLE	Integration	Martin-Marietta	ETL
	Terrain Data Base	Engineering Topographic Laboratory (ETL)	
FLEET COMMAND CENTER BATTLE MANAGEMENT	Force Requirements Expert System (FRESH)	Texas Instruments	NAVELEX
	Spatial Data Management System (SDMS/VIEW)	CCA	NAVELEX
	Combat Action Team (CAT)	Carnegie-Mellon	NAVELEX
	Fleet Command Center Battle Management System	Naval Ocean Systems Command (NOSC)	
	Combat Action Team System	NOSC	
PILOT'S ASSOCIATE	Functional Mission Definition	Perceptronics	DSSW
	Laboratory Definition Study	Draper Labs	DSSW
AIRLAND BATTLE MANAGEMENT	Application Definition	Mitre	DSSW
MULTIPROCESSOR SYSTEM ARCHITECTURES	Tree Machines	Columbia	NAVELEX
	Dataflow Emulation Facility	Massachusetts Institute of Technology	ONR
	Software Workbench	Carnegie-Mellon (CMU)	NAVELEX
	Programmable Systolic Array	Carnegie-Mellon	NAVELEX
	Connection Machine	Thinking Machines	NAVELEX
	Multiprocessor ADA Compiler	Incremental Systems	NAVELEX
NATURAL LANGUAGE UNDERSTANDING	Language Generation	University of Massachusetts	ONR
EXPERT SYSTEM TECHNOLOGY	Reasoning With Uncertainty	General Electric	RADC
SPEECH UNDERSTANDING	Noisy Speech Technology	Texas Instruments	NAVELEX
	Continuous Speech Understanding	Carnegie-Mellon	NAVELEX
	Acoustic Phonetics	Massachusetts Institute of Technology	NAVELEX
	Noisy Speech Front-End	Lincoln Laboratories	ESD
	Speech Co-Articulation	National Bureau of Standards	

PROJECT AWARDS (Continued)

ACTIVITY	PROJECT	PERFORMER	AGENT
COMPUTER VISION	Vision Based Navigation	Univ. of Maryland	ETL
	Terrain Following	Carnegie-Mellon	ETL
	Parallel Algorithms	Carnegie-Mellon	ETL
	Parallel Algorithm Environment	Univ. of Rochester	ETL
	Target Motion Detection	Univ. of Massachusetts	ETL
INFRASTRUCTURE	Adaptive Wafer Scale Integration	Lincoln Labs	ESD
	MOSIS Service	USC/ISI	DSSW
	Compact LISP Machine	Texas Instruments	NAVELEX
	Butterfly Multiprocessor	BBN	DSSW
	Computer Benchmarking	Stanford	NAVELEX
	GaAs Process Evaluation and Design Rules	JPL/Cal. Tech	NASA
MICROELECTRONICS	Low Power, Radiation Hardened GaAs VLSI Pilot Line	Rockwell/Honeywell	AFCMD
	GaAs I/O Package	Mayo Clinic	NOSC
	Integrated Optoelectronics	Rockwell	NOSC
	Hybrid Optical Interconnects	Honeywell	NOSC
	Time-Division Optical Multiplexing	NOSC	
PROGRAM MANAGEMENT SUPPORT	Technical and Management Support	BDM	DSSW
	Technical and Management Support	SAI	DSSW

A large number of projects in all areas of the Strategic Computing Program are currently in procurement, with negotiations on key projects rapidly nearing completion and contract award.

STRATEGIC COMPUTING PROGRAM DARPA CONTRIBUTORS

NAME	TITLE/OFFICE
Dr. Albert E. Brandenstein	Deputy Director, Strategic Technology Office
Dr. Charles Buffalano	Deputy Director for Research, DARPA
Mr. Ray E. Chapman	Former Director, Program Management Office
Dr. Rodger Cliff	Program Manager, Engineering Applications Office
Ms. Lynn Conway	Chief Scientist, Strategic Computing; and Assistant Director, Engineering Applications Office
Dr. Robert S. Cooper	Director, DARPA
Mr. John N. Entzminger	Director, Tactical Technology Office
Dr. Craig I. Fields	Deputy Director, Engineering Applications Office
Mr. James C. Goodwyn	Director, Program Management Office
Mr. John W. Hansen	Deputy Director, Tactical Technology Office
Dr. Robert E. Kahn	Director, Information Processing Techniques Office
Mr. Elias M. Kallis	Program Analyst, Program Management Office
Dr. Sherman Karp	Principle Scientist, Strategic Technology Office
Dr. Clinton W. Kelly, III	Director, Engineering Applications Office
Mr. Paul Losleben	Assistant Director, Automation Technology, Information Processing Techniques Office
Mr. Verne L. Lynn	Deputy Director for Technology, DARPA
LTC Michael E. Montie, USA	Program Manager, Tactical Technology Office
Dr. John A. Neff	Program Manager, Defense Sciences Office
CDR Ronald B. Ohlander, USN	Assistant Director, Computer Science, Information Processing Techniques Office
LTC John Retelle, USAF	Program Manager, Tactical Technology Office
Mr. Sven A. Roosild	Assistant Director, Electronic Sciences, Defense Sciences Office
CDR J. Allen Sears, USN	Program Manager, Information Processing Techniques Office
Dr. Stephen Squires	Assistant Director, Architecture, Information Processing Techniques Office
Dr. Anthony J. Tether	Director, Strategic Technology Office

POINTS OF CONTACT

PROGRAM AREA	NAME	TELEPHONE NUMBER
Autonomous Land Vehicle	Dr. Clinton W. Kelly, III	(202) 694-3622
Fleet Command Center Battle Management	Dr. Albert E. Brandenstein	(202) 694-1702
Pilot's Associate	LTC John Retelle, USAF	(202) 694-3512
AirLand Battle Management	LTC Michael E. Montie, USA	(202) 694-8378
Multiprocessor System Architectures	Dr. Stephen Squires	(202) 694-5927
Speech Understanding	CDR J. Allen Sears, USN	(202) 694-5921
Computer Vision	CDR Ronald B. Ohlander, USN	(202) 694-5051
Infrastructure	Mr. Paul Losleben	(202) 694-5037
Microelectronics	Dr. John A. Neff	(202) 694-5800
Natural Language Understanding	CDR Ronald B. Ohlander, USN	(202) 694 5051
Expert System Technology	CDR J. Allen Sears, USN	(202) 694-5921
Carrier Battle Group Battle Management	CDR J. Allen Sears, USN	(202) 694-5921

Additional copies of this Annual Report are available upon written request to the:

Defense Advanced Research Projects Agency
1400 Wilson Boulevard
Arlington, VA 22209-2308

Attn: Strategic Computing Annual Report

Order Code IB85105

SUPERCOMPUTERS AND ARTIFICIAL INTELLIGENCE: RECENT FEDERAL INITIATIVES

UPDATED 09/06/85

BY

Nancy R. Miller

Science Policy Research Division

Congressional Research Service

ISSUE DEFINITION

In 1983 and 1984, Congress approved legislation for several new Federal initiatives to promote the development of state-of-the-art supercomputers and artificial intelligence. Two of the major programs are the National Science Foundation's Advanced Scientific Computing Program and the Defense Advanced Research Project Agency's Strategic Computing Initiative. Initially, the policy debate dwelled on whether the Government needed to assume a larger role in efforts not only to extend the limits of existing technology, but also to respond to challenges from foreign competitors, in particular Japan. Today, the debate has expanded to include questions on whether these projects have been structured to best serve the Nation's scientific, economic, and national security needs.

BACKGROUND AND POLICY ANALYSIS

Supercomputers can be thought of as the fastest computers at any given time. A computer contains a central processing unit (CPU) which performs the arithmetic and logic operations, and a storage system, or "memory," which contains the data to be processed and the instructions for processing. A program is the sequence of instructions that directs the computer to perform specific operations; a collection of programs is called software. In conventional computers, data and instructions must be passed back and forth sequentially through a link between the CPU and the memory. This computer structure or "architecture" -- conceived by John von Neumann more than 30 years ago -- can lead to a bottleneck in processing speed because it forces serial (one unit at a time) rather than parallel (multiple units at a time) operations.

Dramatic improvements in computing capabilities have occurred in the last two decades. In 1966, the fastest computers could perform one million calculations per second; today's state-of-the-art supercomputers can now exceed processing rates of 200 million operations per second. The substantial increases in speed have been made possible by dramatic advances in miniaturization of microelectronic circuits which increases both the number of operations per second and the available memory storage. In addition, new concepts in computer architecture enable current supercomputers to carry out many similar operations in a minimally parallel manner thereby permitting faster processing rates.

Software technology generally has not kept pace with the advances in hardware technology. In the future, the increasing complexities associated with advanced parallel computer systems could make advances in this area more difficult.

For more than 20 years, supercomputers have played a critical role in scientific and engineering advances. Many computer scientists contend, however, that existing supercomputer technology has reached its limits available under current theory and that further advances will require significant research and development (R&D). Extending limits imposed by existing technology not only will expand applications of supercomputers in scientific and engineering experiments, but also will establish a technological base for future "symbolic" computers capable of artificial intelligence.

Numerical Supercomputers

Numerical supercomputers are large-scale "number crunchers" used for modelling or simulating scientific and engineering problems. Three U.S. companies offer state-of-the-art supercomputers. Cray Research Corporation has 88 machines installed both here and abroad and controls nearly two-thirds of the world market. The second dominant U.S. manufacturer is Control Data Corporation, followed by Denelcor Inc. A new U.S. player is ETA systems, an independent company established by Control Data Corporation (CDC) in 1983, which is developing the successor to CDC's state-of-the-art supercomputer.

The United States generally is recognized as the leader in superspeed computing technology as applied to: the design and simulation of very large-scale integrated circuits; design of nuclear weapons; design and analysis of nuclear reactors; fusion energy research; intelligence applications; aerodynamic design and evaluation; meteorological forecasting; and petroleum exploration. Two Japanese manufacturers, however, recently have produced supercomputers that compare favorably with present generation American systems.

Such challenges have heightened the sense of need among U.S. supercomputer users for increased R&D to produce future generations of supercomputers. Specifically, research requiring attention includes: greater miniaturization of integrated circuit devices; high level parallel computer architectures; and associated software. Providing software capable of managing parallel computation may prove to be the most formidable barrier to future technological advances in superspeed computing. According to a paper in the report by the Panel on Large Scale Computing in Science and Engineering (or the Lax Panel):

> ...it is extremely difficult to plan software algorithms and numerical methods that will take full advantage of this parallelism ...Even with the currently existing minimal parallelism...we do not yet know how to produce near optimal software. It will be even more difficult for the coming generations of supercomputers.

Symbolic Computers

In addition to solving computationally complex problems, future supercomputer developments are likely to advance research in symbolic computing. Symbolic computers can be defined as computers that manipulate non-numerical symbols to carry out tasks which imitate human intelligence. Since the 1950s, attempts have been made to empower digital computers--which had been invented to perform arithmetic and logical operations on discrete quantified data -- with artificial intelligence.

Today, the few practical applications of artificial intelligence have been developed on computer systems which are significantly less powerful than the current generation of numerical supercomputers. Examples of these systems include: "expert" or knowledge-based systems, which give advice and make decisions in highly specific subject areas such as medical diagnosis; and natural language systems.

The programs of these knowledge-based systems solve problems using

heuristic methods or "rules of thumb" rather than algorithms in order to reach a conclusion. "Knowledge" generally is represented by "if-then" rules that suggest alternative courses of action, or patterns which indicate how various symbols are related. Although symbolic computers require different software than numerical supercomputers, both are likely to require high-level parallel architectures and advanced component performance to speed logical inferences and to accommodate vast increases in memory size.

The Defense Advanced Research Projects Agency (DARPA) of the Defense Department is attempting to develop advanced symbolic computers in their Strategic Computing Initiative (SCI). Described by DARPA officials as "superintelligent computers," the goal is to create superspeed computers with human-like capabilities to monitor, reason, plan, and navigate for the following applications:

o Autonomous systems or advanced robotic devices that can sense and interpret their environment and perform a variety of tasks;

o Battlefield management systems to assist commanders in determining courses of military action; and

o Collaborative systems such as a pilot's assistant that can respond to spoken commands.

Symbolic computers capable of these applications will require computational speeds of billions of instructions per second as well as breakthroughs in software.

In addition, advanced symbolic computers would be needed if the Nation decided to build a ballistic missile defense system such as that being studied under the Defense Department's proposed Strategic Defense Initiative (SDI) or "Star Wars" program. One scenario for an SDI system involves four tiers of ground and space-based weapons (see IB81123: "Star Wars"). According to a recent news article, SDI officials have estimated that the software for this application could require nearly 10 million to 100 million lines of software code, or more than 100 times as much as the largest existing computer program.

FOREIGN INITIATIVES

The most highly publicized foreign efforts in the area of supercomputer and artificial intelligence R&D are two Japanese programs -- the National Superspeed Computer Project and the Fifth Generation Computer Project. These targeted national programs are designed not only to meet future domestic needs in large-scale computing and AI, but also to enable Japan to gain access to international markets for these products. Some observers view these Japanese projects as a direct challenge to U.S. leadership in these areas. Other foreign initiatives which are focusing on related areas of information technology include: Great Britain's Alvey Program for Advanced Information Technology, and the European Community's (EC) European Program for Research and Development (ESPRIT).

The Japanese National Superspeed Computer Project is a ten-year program designed to build a new generation supercomputer. The program is a cooperative effort between the Ministry of International Trade and Industry (MITI) and six major Japanese computer vendors. By the end of the project in

1990, MITI will have been obligated, under current plans, to supply about one-half of the total $200 million funding.

Work on Japan's ten-year Fifth Generation Computer Project began in 1982 under the direction of the Institute for New Generation Computer Technology (ICOT), formed by MITI. The goal of the Fifth Generation program is to exploit the hardware developments in the Superspeed Computer Project to create supercomputers capable of AI applications. Like the supercomputer project, this program is a joint endeavor between MITI and private companies. Although total funding for the life of the project is not publicly stated, some sources estimate that up to one billion dollars may be spent. The Japanese Government has supplied seed money and staffing; from 1982 to 1984, the government provided approximately $40 million in funding.

Last year, a panel of U.S. experts assessed these Japanese efforts as part of a research briefing report under the supervision of the Committee on Science, Engineering, and Public Policy (COSEPUP), a joint committee of the National Academies of Sciences and Engineering and the Institute of Medicine. In their study, Report of the Research Briefing Panel on Computer Architecture, the panel stated that Japanese manufacturers have introduced supercomputers that have attained parity with state-of-the-art machines in the United States. In the area of highly parallel supercomputers, however, the panel concluded:

> The United States can expect to retain world leadership in this area, provided that agency support focuses on the most robust new designs and that proper bridges are built between university-centered innovation and the organizational and manufacturing capabilities of the U.S. computer industry.

After examining the Fifth Generation Program, the panel suggested that the ambitious goals set for it are not matched by any entirely persuasive technical concept of how they can be reached. The panel reported, however, that, even if not completely successful, the project will generate significant new engineering, architectural, and software strengths for Japan.

In 1983, the British Government announced the five-year Advanced Information Technology program, which concentrates on AI as well as very large-scale integrated circuits, software engineering, and man/machine interfaces. The government will provide $310 million for the program while private industry is expected to supply an additional $230 million. In analyzing the British effort, the COSEPUP panel concluded that the United States should monitor this potentially significant national effort.

Last year, the EC authorized one billion dollars for the 10-year ESPRIT project. Research will emphasize. advanced microelectronics, software technology, advanced information processing, office automation, and computer integrated manufacturing. The project is a cooperative effort between private companies, universities, and research institutes. The EC will provide $126 million in matching grants in 1985 and $150 million in 1986 for research projects. According to the COSEPUP panel, the ESPRIT program does not appear to confront the United States with competition as severe as that seen in Japan; this effort, however, needs to be closely monitored.

U.S. GOVERNMENT INITIATIVES

The U.S. Government has supported the domestic computer industry since its inception through R&D funding and procurement. The government has stimulated the supercomputer industry directly by purchasing or leasing more than 50 percent of the present generation machines in service in the United States. In the area of symbolic computing, for the past two decades DARPA has funded a large proportion of AI research.

Fiscal years 1984 and 1985 marked the beginning of renewed Federal support for supercomputers and AI. Congress approved funding in the budgets for the National Science Foundation (NSF) to begin establishing new supercomputer centers at universities for the scientific and engineering research communities. Congress also approved funding for DARPA for the Strategic Computing Initiative in artificial intelligence. In addition, three other agencies -- the Department of Energy, the National Aeronautics and Space Administration, and the National Security Agency -- have received appropriations for expanded efforts on behalf of supercomputers.

National Science Foundation Program

The NSF program to establish Advanced Scientific Computing Centers comes in response to claims by several expert panels of the need for greater access by researchers and students to state-of-the-art supercomputers. In 1982 and 1983, three committees -- the Lax Panel, the NSF Working Group on Computers in Research, and the Federal Coordinating Council for Science, Engineering, and Technology (FCCSET) -- concluded that most researchers and students in science and engineering do not have access to supercomputers, and as a result, are unable to address many current research problems. The Lax Panel recommended a federally coordinated national program to increase access for researchers to supercomputer facilities. The NSF Group suggested establishing ten university supercomputer centers and the telecommunications networks to link them to other schools. The FCCSET committee also recommended that Federal agencies design and establish new supercomputer centers and associated networks as needed to meet long-term objectives.

In August 1984, the Office of Advanced Scientific Computing initiated the NSF program by purchasing time from three existing supercomputer centers to provide access to researchers. In FY85, additional "interim centers" also will supply additional time. The interim centers will be phased out as the National Advanced Scientific Computing Centers begin operating.

In February 1985, NSF announced the selection of four institutions that will receive approximately $200 million over the next five years to establish and operate these national centers. The institutions -- University of California at San Diego, University of Illinois at Urbana-Champaign, Princeton University, and Cornell University -- each will receive $7 million to $13 million per year over the grant period. Each award will have a cost-sharing provision in which the states, industries, and institutions will contribute an amount that will approximately double the NSF award.

In late 1985 or early 1986, the centers should be available for use by the scientific and engineering research communities. Plans call for the supercomputer centers to be connected via a nationwide, high-speed network to allow researchers to communicate with the centers from any location. In addition to providing high quality advanced computing systems for researchers, the centers will educate students and researchers in the use of supercomputers.

DARPA's Strategic Computing Initiative

In 1983, DARPA launched the Strategic Computing Initiative. DARPA plans to request approximately $600 million in funding for the first five years of the program. According to a recent report by the Office of Technology Assessment (OTA), Information Technology R&D: Critical Trends and Issues, Congressional approval of SCI funding would increase the annual Federal funding of R&D in AI and related hardware from $30 million to approximately $120 million. Research in the areas of speech, vision, natural language, and expert systems coincides with the main elements of Japan's Fifth Generation Project.

As described in the 1983 SCI planning document, industry will develop the military applications, drawing on results of university research. Advanced computer architectures primarily will be produced in joint endeavors between universities and industry; most of the hardware and software projects will be selected on a competitive basis. In the February 1985 annual report on SCI, DARPA reported that 48.4 percent of the FY84 budget was allocated to industry, and 43 percent was distributed to universities.

Initiatives by Other Federal Agencies

In 1983 and 1984, three other Federal agencies received funding for new supercomputer initiatives. The National Aeronautics and Space Administration has begun work on an advanced supercomputer for designing and testing commercial and military aircraft. NASA estimates the cost of the numerical aerodynamic simulator project to be approximately $120 million, from FY84 to FY88.

Last year, the Department of Energy received $7 million to expand access to supercomputers for energy research scientists. Congress also approved $7 million in the energy and water development appropriations for FY85 to establish a supercomputer center at Florida State University.

In November 1984, the National Security Agency announced the establishment of a Supercomputer Research Center. The $12 million facility will be built at the Maryland Science and Technology Center by the Institute for Defense Analyses, the contractor for the project. The goal of this endeavor is to design and produce a next generation supercomputer.

POLICY DEBATE

In 1983 and 1984, Congress considered several measures designed to increase the Federal role in supercomputers and artificial intelligence. During that time, the policy debate concerned whether the government needed to assume a larger role in efforts not only to extend the limits of existing technology, but also to respond to challenges from foreign competitors, in particular Japan. Specifically, given the limited market for supercomputers along with the uncertainties of a high-risk market for both supercomputers and AI, would the private sector support the R&D initiatives necessary to make advances in these areas? If not, should the Federal government assume a larger role to ensure continued U.S. leadership in technologies that many consider crucial for scientific advances and national security?

Congress responded to these initial concerns by approving funds to begin

work on several Federal agency initiatives. Today the focus of the debate has expanded to include whether these efforts have been structured to best serve the Nation's scientific, economic, and national security needs.

To date, the major controversy pertains to DARPA's Strategic Computing Initiative, the program which has received the largest amount of funding. Part of the debate concerns the feasibility of achieving truly significant advances in AI that will be necessary to reach the stated goals set forth in the SCI plan. Some experts claim that recent advances in natural language processing, vision systems, and in particular expert systems, are the beginning of a new age in artificial intelligence. Further, some analysts predict that AI could become a multibillion annual business within a decade.

Conversely, other experts assert that current systems are based on 20-year-old programming techniques that have become practical only as the price of computer power has declined. They claim that, despite the successes of several limited-domain expert systems, it will be much more difficult, if not impossible, to develop computer systems capable of true "intelligence" and common sense reasoning. Further, in a recent article, the director of DARPA acknowledges that the scientific ambitiousness of SCI is even greater than the Manhattan Project which developed the atomic bomb. In comparison, he states that the Manhattan Project was a fairly easy job.

In addition to questions about the feasibility of attaining the stated goals of the program, some computer experts question the trend toward use of AI techniques in battle. In a 1984 report, Strategic Computing: An Assessment, the Computer Professionals for Social Responsibility (CPSR) state that increased reliance on AI and automated decision-making in critical military situations leads in a dangerous direction. Like all computers, AI systems only will perform what they are programmed to do and may act inappropriately in unanticipated situations.

That CPSR report also expressed concerned that the technologies of the SCI could be used in a ballistic missile defense system. In discussing the need for autonomous systems in unpredictable military situations, the 1983 SCI planning document states:

> An extremely stressing example of such a case is the projected defense against strategic nuclear missiles, where systems must act so rapidly that it is likely that almost complete reliance will have to be placed on automated systems.

In response to the debate about reliance on AI systems in battle, DARPA claims that, since no one yet knows what a "Star Wars" system would entail if a decision were made to employ one, it is impossible to predict the need for autonomous control of the system. According to a recent statement by the director of DARPA, "Any applications of the technology later evaluated and proposed by the Defense Department for use in actual weapons systems will . . . be subject to the usual administrative processes under the control and guidance of elected public officials." In addition, the DARPA director asserts that the SCI will develop a new generation of machine intelligence technology and prototype applications only to the point of feasibly demonstrating basic results.

Another issue arises from the large amount of funding intended for SCI; such a growth in funding significantly increases DARPA's dominance over the

direction of the Nation's advanced computer research efforts. Computer experts agree that since its inception in 1958, DARPA has been instrumental in advancing computer technology in the United States. Many of the computer technologies funded by DARPA -- such as timesharing, networking, and computer graphics -- have evolved into commercially successful applications.

For the past 20 years, DARPA also has been the major U.S. funding source for AI. According to the OTA report, the SCI program will quadruple the annual Federal funding of R&D in AI and related hardware; by contrast, in recent years, the NSF budget for AI projects has been $5 million to $6 million annually. Expanded funding for SCI not only solidifies DARPA's position in directing AI research priorities, but also has raised issues about the imbalance in the military/civilian funding of this area. These issues include:

-- the emphasis on developing computer technologies with specifically military applications;

-- the skewing of U.S. computer research away from more commercial pursuits;

-- the likelihood of competition among military, commercial, and academic communities for a limited number of qualified AI researchers;

-- the possibility that SCI research will be classified due to the military nature of the program; and

-- the possibility that SCI research will not produce "dual-use" technologies, particularly with the military applications of the program.

Supporters of the SCI program and DARPA officials contend that these concerns are unfounded. Based on DARPA's past successes in computer research, they assert that DARPA is the appropriate Federal entity to administer a large-scale program in AI. In addition, they argue that the SCI will ensure U.S. dominance in computer technology and strengthen the Nation's technological and economic base. According to the SCI planning document:

> The Strategic Computing Program promises the production of machine intelligence technology that will enable yet another major cycle of new economic activity in the computer and electronics industry. If the United States aggressively competes to develop these systems, it will gain access to enormous new commercial markets . . . Spinoffs from a successful Strategic Computing Program will surge into our industrial community.

DARPA officials also assert that the bulk of SCI's basic research will be performed by universities, while most of the applied research will be conducted by industry. Although some of the applied research might be classified, the generic research will remain unclassified. In addition, DARPA admits that the SCI may put a short-term strain on the availability of AI researchers; the agency claims, however, that over time the expanded funding will increase the total base of skilled workers. Finally, DARPA officials deny that there has been a shift in emphasis toward military

applications.

Some AI researchers contend that a greater percentage of the funding for AI research should come from non-military sources. To resolve the imbalance between military and civilian funding, some observers have called for a "civilian DARPA" or expanded activity by an existing Federal agency such as NSF. To date, NSF's primary response to the need for greater activity in the area of advanced computer technology has been its program to establish supercomputer centers at universities. These centers likely will educate computer science students and researchers in the use of supercomputers. The NSF program, however, primarily has been designed to provide high-speed computing tools for the general scientific and engineering community rather than to advance research in supercomputers and AI.

Conversely, other experts claim that, over the long term, the source of funding does not determine the potential utility of a program's research results. They assert that no other Federal agency could be as efficient and productive as DARPA in managing a large-scale effort in advanced computer technology. Considering the importance of supercomputers and AI to our technological and economic base, Congress may want to examine and monitor these new initiatives to ensure that they best serve the needs of the Nation.

LEGISLATION

Major Funding Initiatives: 98th Congress

The major funding initiatives in the areas of supercomputers and AI are NSF's Advanced Scientific Computing Program and DARPA's Strategic Computing Initiative. In FY84, DARPA received $50 million in appropriations legislation (P.L. 98-212) for the SCI; in FY85, DARPA program received $72 million in the continuing appropriations legislation (P.L. 98-473). NSF received $40 million for the supercomputer center program in the FY85 appropriations for HUD-Independent Agencies (P.L. 98-371).

For FY86, DARPA has requested $142 million for the SCI, while NSF has asked for $46.2 million for Advanced Scientific Computing.

National Science Foundation

H.R. 3038 (Boland)

Appropriations for HUD-Independent Agencies for FY86. Within Title II, Independent Agencies, recommendation of $46.2 million for NSF's Advanced Scientific Computing program. Reported to House by Committee on Appropriations (H.Rept. 99-212) on July 18, 1985. House passed on July 25, 1985. Senate Committee on Appropriations ordered H.R. 3038 to be reported with amendments favorably on July 31, 1985.

H.R. 1210 (Fuqua)

Authorizes NSF appropriations for FY86. Within Section 2, request for $46.2 million for Advanced Scientific Computing. Reported to House by Committee on Science and Technology (H.Rept. 99-44) on Apr. 16, 1985. Passed House Apr. 17, 1985.

S. 801 (Hatch)

Authorizes NSF appropriations for FY86. Within Sec. 2, request for $46.2 million for Advanced Scientific Computing. Introduced Mar. 28, 1985; reported by Committee on Labor and Human Resources on July 17, 1985.

S. 903 (Danforth)

Authorizes NSF appropriations for FY86. Within Sec. 2, request for $46.2 million for Advanced Scientific Computing. Reported by Committee on Commerce, Science, and Transportation (S.Rept. 99-27) on Apr. 15, 1985.

Department of Defense

S. 1029 (Goldwater)

Authorizes Armed Forces appropriations for FY86. Within Title II, RDT&E, request for $142 million for Strategic Computing Initiative (SCI). Reported to Senate by Committee on Armed Services (S.Rept. 99-41) with request for $132 million for SCI on Apr. 29, 1985.

S. 1160 (Goldwater)

Authorizes Armed Forces appropriations for FY86. Reported by Committee on Armed Services as an original measure in lieu of S. 1029 on May 16, 1985 (without written report). Senate agreed to Amendment No. 182 which restored $42 million to SCI that had been cut by the Committee on Armed Services on May 21, 1985. Senate passed S. 1160 with amendments on June 5, 1985 with request of $142 million for SCI. Conference reports (H.Rept. 99-235 and S.Rept. 99-118) filed on July 29, 1985, with request of $142 million for SCI. Senate agreed to conference report on July 30, 1985.

H.R. 1872 (Aspin)

Authorizes Armed Forces appropriations for FY86. Within RDT&E for Defense Agencies, request for $142 million for the Strategic Computing Initiative. Reported to House by Committee on Armed Services (H.Rept. 99-81) on May 10, 1985. House passed on June 27, 1985. House incorporated H.R. 1872 in S. 1160 as an amendment on June 27, 1985.

HEARINGS

U.S. Congress. House. Committee on Science and Technology. Subcommittee on Energy Development and Applications and Subcommittee on Energy Research and Production. Computers and their role in energy research: current status and future needs. Hearings, 98th Congress, 1st session. Washington, U.S. Govt. Print. Off., 1983. 413 p.

U.S. Congress. House. Committee on Science and Technology. Supercomputers. Hearings, 98th Congress, 1st session. Washington, U.S. Govt. Print. Off., 1983. 460 p.

ADDITIONAL REFERENCE SOURCES

Alexander, Tom. Why computers can't outthink the experts. Fortune, v. 110, Aug. 20, 1984: 105-118.

Artificial intelligence: can parts and programs make a brain? Science 85, v. 6, Mar. 1985: 38-51.

Computer Professionals for Social Responsibility. Strategic computing: an assessment. Palo Alto, Cal., CPSR, 1984. 4 p. (Reprinted in Bulletin of the atomic scientists, v. 40, Dec. 1984: 11-15).

Davis, Dwight B. Assessing the strategic computing initiative. High technology, v. 5, Apr. 1985: 41-49.

Edwards, Paul N. Are 'intelligent weapons' feasible? The nation, v. 240, Feb. 2, 1985: 110-112.

The race for computer supremacy: who's ahead? New york times, Oct. 23, 1984: C1, C8.

Sanger, David E. The surge in supercomputers. New york times, Mar. 1, 1985: D1, D16.

Supercomputers come out into the world. Economist, v. 292, Aug.11, 1984: 77-80.

U.S. Library of Congress. Congressional Research Service. "Star wars": antisatellites and space-based BMD. Issue Brief No. IB81123, by Marcia S. Smith, Feb. 20, 1985 (continuously updated). Washington, 1985.

Waldrop, M. Mitchell. The fifth generation: taking stock. Science, v. 226, Nov. 30, 1984: 1061-1063.

AN ANNOTATED BIBLIOGRAPHY OF LITERATURE ON SUPERCOMPUTERS

J. C. BROWNE *
JOHN FEO **
PATRICIA ROE ***

DRAFT 1.0

MAY 1985

* Professor, Computer Science Department, The University of Texas at Austin

** Research Assistant, Computer Science Department, The University of Texas at Austin

*** Research Assistant, IC^2 Institute, The University of Texas at Austin

1. INTRODUCTION

This annotated bibliography is an attempt to compile a focused summary of the current literature on "Supercomputers" and their applications. Selection of articles to include and annotate was subjective. The criterion for selection had as one common base the concept of the most powerful computers which can be built and applied in a given problem environment.

The current listings are known to be deficient in many areas. Many citations are incomplete - missing authors, publication dates, etc. These will be supplied in a later draft. Work is continuing. The compilers will be very grateful for any additional citations to which the readers of this preliminary draft can call their attention.

We are grateful to Cray Research for providing us with their literature survey. Support for preparation of this bibliography has been provided by the RGK Foundation of Austin, Texas, the IC^2 Institute at the University of Texas at Austin and the Westinghouse Electric Corporation.

2. APPLICATIONS

[1] Adelantado,M., Comte,D., Siron,P. and Syre,J.C.
An MIMD Supercomputer System for Large-Scale Numerical Applications.
In *Proc. of the IFIP 9th World Computer Congress*, Pages 821-826. North Holland, Amsterdam, Netherlands, Sept, 1983.

The research described uses a global approach to multiprocessing in which software and hardware studies have been conducted together. Several MIMD computer organizations and their corresponding languages were selected. Simulations were performed and now with results appearing, both language and architecture are being refined to satisfy the original goals. This study was done in the context of the French supercomputer project, MARIANE.

[2] Anderson,W., Brice,R. and Alexander,W.
Configuring a Supercomputer for an Interactive Scientific Workload.
Technical Report LA-UR-82-1572, Los Alamos National Labs., 1982.

A detailed, validated simulation model of an existing Cray-1 running under an interactive operating system was used to investigate configurations of a new supercomputer recently announced by the same vendor. The goal was to determine the optimum configuration for a known interactive scientific workload. Questions considered included how much main memory would be needed and whether to acquire an optional fast swapping device.

[3] Ballhaus, W. F. Jr.
Supercomputing in Aerodynamics,
Frontiers of Supercomputing. University of California Press, .

[4] Ballhaus,W.F.
Computational Aerodynamics and Supercomputers.
Technical Report NASA-TM-85887, NASA, 1984.

Some of the progress in computational aerodynamics over the last decade is reviewed. The Numerical Aerodynamic Simulation Program objectives, computational goals, and implementation plans are described.

[5] Bardon, M. and Curtis, K. K.
A National Computing Environment for Academic Research.
Technical Report , National Science Foundation, , 1983.

This report is the genesis of the Advanced Scientific Computing Program now put in place by the National Science Foundation.

[6] Barlow,R.H., Evans,D.J. and Shanehchi,J.
Comparative Study of the Exploitation of Different Levels of Parallelism on Different Parallel Architectures.
In *Proc. of the 1982 Inter. Conf. on Parallel Processing*, Pages 34-40. IEEE Computer Soc. Press, Silver Springs, Md., Aug, 1982.
The authors identify three levels of parallelism obtainable from sequential solutions for locating the eigenvalues of real symmetric tridiagonal matrices based on the bisection algorithm coupled with Strum sequence evaluation. The implementation of these three levels are compared for the CRAY-I, the ICL DAP and an asynchronous, shared-memory multiprocessor.

[7] Bersohn,M.
Supercomputers and the Problem of Organic Synthesis,
ACS Symposium Series, Pages 109-116. American Chemical Society, Washington, D.C., 1981.
Annotation - TBD.

[8] Bishop,M., Ceperley,D., Frisch,H.L. and Kalos,M.H.
The Need for Supercomputers in Time-Dependent Polymer Simulations,
ACS Symposium Series, Pages 135-142. American Chemical Society, Washington, D.C., 1981.
Annotation - TBD.

[9] Boris, J. P. and Winsor, N. K.
Vectorized Computation of Reactive Flow,
Parallel Computations, Vol. 1, Computational Techniques, Pages 173-214. Academic Press, 1982.

[10] Bowler, K. C. and Pauly, G. S.
Molecular Dynamics and Monte Carlo Simulations in Solid State in Elementary Particle Physics.
In *Proceedings of the IEEE 72*, Pages 42-55. IEEE, 1984.
This article describes applications of a ICLDAP to a number of problems from dynamics, lattice gauge theory and the Izing model. The article discusses the techniques of application of a square array processor and gives some of the mathematical models and results from their simulations.

[11] Buzbee, B. L.
Gaining Insight from Supercomputing.
In *Proceedings of the IEEE*, Pages 19-21. IEEE, 1984.
This paper gives examples of the application of very high performance computing to obtain insight into qualitative behavior of complex systems.

[12] Byrne, G.W.
A View of Supercomputing,
Frontiers of Supercomputing. University of California Press, .

This article describes the several applications of very large scale computation in the Exxon Company. These applications include reservoir simulation, transport modeling and others.

[13] Calahan,D.A.
A Block-Oriented Sparse Equation Solver for the CRAY-I.
In *Proc. of the 1979 Inter. Conf. on Parallel Processing*, Pages 116-123. IEEE Computer Soc. Press, New York, N.Y., Aug, 1979.
Annotation - TBD.

[14] Collins,J.F. and Coulson,A.F.
Applications of Parallel Processing Algorithms for DNA Sequence Analysis.
*Nucleic Acids Research*12(1):181-192, Jan, 1984.
The authors describe parallel processing algorithms for the DNA sequence analysis and the implementation of these algorithms on the ICL DAP. It is noted, that since parallel processing is so efficient a wider range of comparison problems may be considered than has heretofore been the case.

[15] Computational Mechanics Committee.
Computational Mechanics.
Technical Report , National Academy of Science Press,
.
This report describes the applications of computational mechanics in simulation of a wide variety of areas and disciplines. It called attention to the need for applications of very high performance computing in order to lower the cost and improve the quality of designs.

[16] Cray Staff.
Cosmic Modeling.
Cray Channels:65-69, 1984.
This article describes applications of very large scale computation to the numerical modeling of the movement of neutrinos in the universe. They call attention, in particular, to the need for very powerful computation to increase the reality level of research modeling.

[17] Cromie, W. J. and Edson, L.
Before They Can Speak, They Must Know.
*Mosaic*15:28-35, 1984.
This article describes knowledge based systems and how they will help tomorrow's computers interact with people.

[18] DARPA.
Strategic Computing.
Technical Report , DARPA.
.
This report describes the DARPA strategic computing program. The applications of very high performance, particularly intelligent computing

in military and defense applications. It is a thoughtful document, one which merits the attention of persons interested in the defense oriented applications of large scale computing.

[19] Davis,C.G. and Couch,R.L.
Ballistic Missile Defense: A Supercomputer Challenge.
*Computer*13(11):37-46, Nov, 1980.
The Ballistic Missile Defense Problem is presented and its need for exceptionally high execution rates and high reliability shown. The supercomputer research and development that BMD inspired in the 1960s and 1970s and the challenges it poses for the supercomputers of the 1980s are delineated.

[20] Demos, G., Brown, M. D. and Weinberg, R. A.
Digital Scene Simulation: The Synergy of Computer Technology and Human Creativity.
In *Proceedings of the IEEE 72*, Pages 22-31. IEEE, 1984.
This article is a semi-popular article describing the techniques used by digital productions for generating digital simulations of scenes for application in motion pictures. The work is done on a very high performance computer (a Cray XMP).

[21] Dennis,J.B., Gao,G.R. and Todd,K.W.
Modeling the Weather with a Data Flow Supercomputer.
*IEEE Transactions on Computers*C-33(7):592-603, July, 1984.
The performance improvement that dataflow computers offer is shown for a NASA benchmark program that implements a global weather model. The authors specify the capacities of hardware units and the number of each type of unit required to achieve a twenty-fold improvement in performance for the weather simulation application.

[22] Department of Energy.
The Role of Supercomputers in Energy Research Programs.
Department of Energy, 1985.
This report from the Department of Energy describes the applications of very large scale computing in the research programs of the Department of Energy. It is the result of a meeting held to prepare a report on this subject.

[23] Doyle, J.
What Should A.I. Want from the Supercomputers.
Technical Report , CMU,
CS-83-160.
This article will also appear in A.I. magazine. This article examines the requirements for parallel structuring of prologue. It describes some of the needs for more radical, non-traditional architectures to support A.I. type

computations.

[24] Dubois, P. F.
Swimming Upstream: Calculating Table Lookups and Piecewise Functions.
Parallel Computations, Vol. 1, Computational Techniques, Pages 129-151. Academic Press, 1982.

[25] Egan,J.T. and MacElroy,R.D.
Calculating the Electrostatic Potential of Molecular Models with Separate Evaluations by Conventional, Vector and Array Processors.
*Journal of Computational Chemistry*5(1):52-81, Jan, 1984.
Annotation - TBD.

[26] Elson,B.M.
Supercomputer Sought to Aid Design.
*Aviation Week*108:123-126, Apr 24, 1978.
Annotation - TBD.

[27] Evans, D. J., Hanley, H. J. M., and Hess, S.
Non-Newtonian Phenomena in Simple Fluids.
Physics Today 37:26-35, 1984.
This article gives the results of formulations of non-equilibrium fluids on a very high performance computer.

[28] Fichtner, W., Nagel, L. W., Penumalli, B. R., Peterson, W. P., D'arcy, J. L.
The Impact of Supercomputers on IC Technology Development and Design.
In *Proceedings of IEEE 72*, Pages 96-112. IEEE, 1984.
This paper reviews some of the computer-aided design tools that are vey high requirements for computational power. They include process simulation, device simulation and circuit simulation. The paper also reviews the impact of very high performance computing on these programs at AT&T Bell, Laboratories.

[29] Fichtner, W., Nagel, L.W., Penumalli, B.R., Peterson, W. P., D'arcy, J.L.
The Impact of Supercomputers on IC Technology Development and Design.
In *Proceedings of the IEEE 72*, Pages 96-112. IEEE, 1984.
This paper reviews some of the computer-aided design tools that have very high requirements for computational power. They include process simulation, device simulation and circuit simulation. The paper also reviews the impact of very high performance computing on these programs at AT&T Bell Laboratories.

[30] Fox, G. C., and Otto, S. W.
Algorithms for Concurrent Processors.
Physics Today 37:53-65, 1984.
This article describes applications of the Cal Tech Cosmic Cube machine. Cosmic Cube has a hyper cube connectivity between the computers in an

array. This article describes several successful applications of computers in problems with regular topological properties.

[31] Fuss, D. and Tull, C. G.
Centralized Supercomputer Support for Magnetic Fusion Energy Research.
In *Proceedings of the IEEE 72*, Pages 32-41. IEEE, 1984.
This article describes the work done at the magnetic fusion energy computation center. It describes the applications, role and structure of the NMFECC supporting research in fusion physics.

[32] Garfinkel, D.
We Could Wire Up an Intelligent Artificial Pancreas- But What Would We Tell It To Do?
Perspectives in Computing 48:10-16, 1984.
This is a very interesting article describing what would have to be done to construct a computer which could model the execution of a pancreas. It explores the benefits which could be obtained and the problems which would have to be solved.

[33] Gloudeman, J. F.
Anticipated Impact of Supercomputers and Finite Element Analysis.
In *Proceedings of the IEEE 72* , Pages 80-89. IEEE, 1984.
This paper describes the status of the application of supercomputers in finite element analysis and the anticipated future impact. It projects the role of very high performance computing in allowing design optimization of non-linear analysis and a closer link between designers and analysts.

[34] Goddard, W. A.
Theoretical Chemistry Comes Alive: Full Partner with Experiment.
Science 27:91-923, 1985.
This article describes the advances in computational techniques in theoretical chemistry in recent years. It gives examples of the predictions which have resulted from these applications. The ability of computational engines to support theoretical chemistry is stressed.

[35] Graedel, T. E., and McGill, R.
Graphical Presentation of Results from Scientific Computer Models.
Science 15:1191-1198, March, 1982.
This article describes the use of graphical modeling to presentation of results from very complex computations.

[36] Guest,M.F. and Wilson,S.
The Use of Vector Processors in Quantum Chemistry,
ACS Symposium Series, Pages 1-37. American Chemical Society, Washington, D.C., 1981.
Annotation - TBD.

[37] Happ,H.H.
Parallel Processing in Power Systems.
In *Proc. of the 7th Power Sys. Comp. Conf.*, Pages 9-16. Westbury House, Guildford, England, July, 1981.
The influence and effects of supercomputers, multiprocessor networks and array processors on power system computing is discussed.

[38] Happ,H.H.
Parallel Processing in Power Systems.
*Inter. Journal of Electr. Power and Energy Systems*4(1):37-41, Jan, 1982.
A brief status report is presented of the emerging computer architectures and their probable influence on power-system computing.

[39] Harlow, F. H.
Superproblems for Supercomputers,
Frontiers of Supercomputing. University of California Press, .
This article characterizes the general property of problems which can effectively utilize supercomputers. Characterization includes, for example, descriptions of systems which have many particles interacting. It includes performance of experiments which involve very complex relationships and mathematical structures. A third is the testing of behavior models. This characterization is a very interesting way to approach the structuring of computational problems.

[40] Heine, R. W.
CAD/CAM: Ford Motor Company Engineering and Manufacturing,
Frontiers of Supercomputing. University of California Press, .
This article describes work at Ford Motor Company in computer aided design and computer aided manufacturing. The focus of the article is on applications which require very large computational capabilities. The fairly detailed article describing many of the applications in the actual engineering and manufacturing process.

[41] Henderson, D.
Computation: The Nexus of Nuclear Weapon Development,
Frontiers of Supercomputing. University of California Press, .
This article describes the applications of very large computers in the design of nuclear weapons. It covers the properties of the codes, the running time, and the general procedure of developing design using these very large scale computers.

[42] Herr, R. A.
Computer Models Gaining on El Nino.
Science:37-38, July, 1984.
This article is a summary of work presented on atmospheric simulations with the purpose of predicting the occurrence of the el nino phenomenon.

[43] Herr, R. A.
Climates Since the Ice Age Began to Melt.
Science 226:326-327, .
This article describes the research on climatic modeling in the archeological past. These models are very complex and require the use of the most powerful possible computers.

[44] Hood, R. T., and Kennedy, K.
Programming Language Support for Supercomputers,
Frontiers of Supercomputing. University of California Press, .
This article deals primarily with current status of techniques for detecting parallel structuring in programs written in Fortran. It also describes a projected environment for the development of vector and parallel programs being constructed by Rice University.

[45] Hoover, W. G.
Computer Simulation of Many Bodied Dynamics.
Physics Today 37:44-50, 1984.
This article describes the methods of simulation of non-equilibrium molecular dynamics. The methods described are explicitly adapted for use on a very large scale computer.

[46] Horton, E. J. and Compton, W. D.
Technological Trends in Automobiles.
Science:587-593, August, 1984.
This article describes many applications of computing to the design and development of automobiles. It points out the requirements for very high performance modeling in many areas of the manufacturing process for automobiles.

[47] Huff, R. W., Dawson, J. M., and Culler, G. J.
Computer Modeling in Plasma Physics on the Parallel-Architecture CHI Computer,
Parallel Computations, Vol. 1, Computational Techniques, Pages 365-396. Academic Press, 1982.

[48] Ishiguro,M., Harada,H., et.al.
Applicability of Vector Processing to Large-Scale Nuclear Code.
Technical Report JAERI-M-82-018, Japan Atomic Energy Research Inst., 1982.
The authors state that to meet the growing trend of computational requirements in JAERI, introduction of a high-speed computer with vector processing faculty is desirable. Problems of vector processors are also discussed from the view points of code performance and ease of use. In particular, vector processing efficiency is assessed with respect to large-scale nuclear codes.

[49] James,R.A.
Simulation of Particle Problems in Astrophysics.
*Computer Physics Communications*26(3-4):423-431, June, 1982.
The author reports that a speed-up of a factor between 7 and 10 in the particle-mesh simulation techniques used in the study of galactic evolution on the CRAY-I and the ICL DAP. The author notes that while a substantial amount of program revision is required, the basic algorithms remain unchanged.

[50] Johnson, O. G.
Three-Dimensional Wave Equation Computations on Vector Computers.
In *Proceedings of the IEEE 72* , Pages 90-95. IEEE, 1984.
This paper describes work on development of computational models of three-dimensional analysis of interest to seismologists. One model is developed in detail. The technical aspects of implementing this algorithm on a vector processor is presented.

[51] Karplus,W.J.
The Impact of New Computer System Architectures on the Simulation of Environmental Systems.
In *Conf. on Modeling, Identification and Control in Environmental Sys.*, Pages 1001-1009. North Holland, Amsterdam, Netherlands, Aug, 1977.
The new class of computer systems introduced in recent years, including supercomputers, array processors, networks of minicomputers and microprocessors and special purpose hybrid systems are surveyed. Their impact upon the simulation of environmental systems characterized by ordinary and partial differential equations is discussed.

[52] Kershaw, D.
Solution of Single Tridiagonal Linear Systems and Vectorization of the ICCG Algorithm on the Cray-1,
Parallel Computations, Vol. 1, Computational Techniques, Pages 85-99. Academic Press, 1982.

[53] Killough,J.E.
The Use of Vector Processors in Reservoir Simulation.
In *Proc. 5th Symp. on Reservoir Simulation*, Pages 21-28. Soc. Pet. Engrs. of AIME, Dallas, Tx., Jan, 1979.
Tests of reservoir simulation models on the IBM 3838 array processor and CRAY-1 computer are described. The formulation of the model and the codifications which were necessary for the vector processors are given. It is shown that vector processors can result in substantial improvements in computation time for the reservoir simulation problem.

[54] Kincaid,D.R. and Oppe,T.C.
ITPACK on Supercomputers.
In *Proc. of the 1982 Inter. Workshop on Numerical Methods*, Pages 151-161.
Springer-Verlag, Berlin-New York, June, 1982.
Annotation - TBD.

[55] Kolata, G.
A Fast Way to Solve Hard Problems.
Science:1379-1380, September, 1984.
This article describes the possible implications of the linear programming algorithm discovered by Karmarker. The article mentions the possibility that a sufficiently powerful parallel computer could use this algorithm do develop results in real time.

[56] Krzeczlowski,A.J., Smith, E.A. and Gethin,T.
Seismic Migration Using the ICL Distributed Array Processor.
*Computer Physics Communications*26(3-4):447-453, June, 1982.
The paper evaluates the DAP for the seismic migration portion of the seismic processing sequence. The performance of the DAP is compared against conventional computers for two different program implementations: the first using vector algorithms and the second reformulating the problem to take full advantage of the parallel architecture.

[57] Kuck,D.J.
ILLIAC IV Software and Application Programming.
*IEEE Transactions on Computers*C-17(8):758-770, Aug, 1968.
The ILLIAC IV operating system, applications programming and high-level language are described. Program examples are included.

[58] Lax, B. D.
Report of the Panel on Large Scale Computing in Science and Engineering.
Technical Report , Lax Panel,
.
This report is a summary of the requirements statement for support of very large scale computing in science and engineering. It is one of the landmark documents in the development of the current programs for access to supercomputers in science and engineering.

[59] Leontief, W., Duchin, F. and Szyld, D. B.
New Approaches in Economic Analysis.
*Science*228:419-422, 1985.
This article describes impact on economic modeling of the availability of very large, very powerful computers. It projects the impact on the relief of this constraint and the effect it will have on the development of the subject.

[60] Lomax, H. and Pulliam, T. H.
A Fully Implicit, Factored Code for Computing Three-Dimensional Flows on the ILLIAC IV,
Parallel Computations, Vol. 1, Computational Techniques, Pages 217-249. Academic Press, 1982.

[61] Lykos,P. and Shavitt,I.
Supercomputers in Chemistry,
ACS Symposium Series. American Chemical Society, Washington, D.C., 1981.
Annotation - TBD.

[62] MacCormack,R.W. and Stevens,K.G.
Fluid Dynamics Applications of the ILLIAC-IV Computer.
In *Proc. of the Comp. Methods and Problems in Aeronautical Fluid Dyn.*, Pages 448-465. Academic Press, London, England, Sept, 1974.
The author reports that the ILLIAC-IV is beginning to yield useful results at computing speeds an order of magnitude faster than the most advanced conventional computer. However, this is still not fast enough to solve all fluid dynamics problems of interest.

[63] Maples,C., Weaver,D., Rathbun,W. and Meng,J.
The Utilization of Parallel Processors in a Data Analysis Environment.
*IEEE Transactions on Nuclear Science*NS-28(5):3880-3888, Oct, 1981.
The Modular Interactive Data Analysis System, MIDAS, is described. In the MIDAS project a multi-processor array consisting of 8 high-level CPUs, 2 pipeline processors and a special high-speed boolean processor is available to each user during analysis.

[64] McCoy,M.G., Mirin,A.A. and Killeen,J.
Vectorized Fokker-Planck Package for the Cray-1.
Technical Report CONF-790902-1, Lawrence Livermore National Labs., 1979.
A program to solve the time-dependent, two-dimensional, nonlinear, multi-species Fokker-Planck equation on the Cray-1 is described. Overall, timing tests show that the vectorized Cray program realizes up to a factor of 12 advantage over an optimized CDC-7600 program and up to a factor of 365 over a non-vectorized version of the same program on the CRAY.

[65] McKerrell,A. and Delves,L.M.
Solution of the Global Element Equations on the ICL DAP.
*ICL Tech. Journal*4(1):50-58, May, 1984.
The Global-Element Method, which is a variant of the finite-element method for the solution of elliptic partial differential equations, is described in this paper. The results of an implementation of a two-dimensional GEM program on a ICL 2980 computer are reported.

[66] McLean,W.J.
Report of the Ad-Hoc Combustion Research Facility Comm. on Computational Resources for Combustion Research.
Technical Report SAND-83-8237, Sandia National Labs.,
1983.
This report looks at CRF computer needs for 1983 to 1988. It is determined that advancing the understanding of the chemical and physical processes in combustion will require a rapidly increasing use of fast supercomputers with large memories.

[67] Nelson, D. B.
Supercomputers and Magnetic Fusion Energy,
Frontiers of Supercomputing. University of California Press, .
This article discusses controlled fusion, why computing is important to controlled fusions and the contributions to research from supercomputers. It gives descriptions of how the research in fusion actually uses supercomputers and the problems and opportunities offered by future supercomputers for fusion research.

[68] Nelson,H.R.
New Vector Super Computers Promote Seismic Advancements.
*World Oil*194(1):155-160, Jan, 1982.
Annotation - TBD.

[69] Newton,A.R. and Pdederson,D.O.
Research in Computer Simulation of Integrated Circuits.
Technical Report AFOSR-82-0021, Air Force Office of Scientific Research,
1983.
The performance of the new LSI simulator CLASSIE running on a Cray-1 is evaluated. CLASSIE is since as an important advance for while there are a large number of powerful design verification tools for IC design at the transistor and logic gate level, there are very few silicon-oriented tools for architectural design and evaluation.

[70] No Author Given.
Supercomputing Research and Development: The LLNL Program and Proposals.
This paper describes LLNL's activities and includes recommendations for the goals of such a program. THis report also includes an assessment of future programmatic requirements in computing, a description of LLNL's computational environment, a synopsis of software development activities, the status of the S-1 supercomputer project and related work in microelectronics.

[71] No Author Given.
Massively Parallel Processor Yields High Speed.
Aviation Week & Space Technology:157, May 28, 1984.

It is reported that Goodyear Aerospace's ultra-high-speed, massively parallel processor is being evaluated at NASA's Goddard Space Flight Center. The article goes on to say that while the MPP was developed to analyze imagery an exploration for other application for which the computer's novel architecture will enable it to match or better the performance of general purpose supercomputers will be undertaken.

[72] No Author Given.
Systems & Peripherals: MPP Processes Satellite Data.
Computerworld:99, Feb 13, 1984.
The article reports that Goodyear Aerospace is testing its Massively Parallel Processor at NASA's Goddard Space Flight Center. The MPP is described and its is noted that the MPP outperforms pipelined processors on fixed-point calculations but not on large floating-point applications.

[73] No Author Given.
The National Security Aspects of Supercomputer R & D.
Computer Age:1,4, Dec 19, 1983.
The report states that the most promising applications of supercomputers are in weapons design, weapons effect studies, intelligence gathering and processing, aircraft and automobile design, bio-technology, petroleum engineering and semiconductor design. It then goes on to argue that the US government ought to provide universities with access to supercomputer centers so as to train scientists and engineers in their use.

[74] Norrie, C.
Supercomputers for Superproblems: An Architectural Introduction.
Computer:62-74, March, 1984.
This article discusses supercomputer operating requirements with emphasis on efficient handling of matrices.

[75] Orme,M.
The Ultimate in CAD Tools.
*Technology*8(12):24-25, Mar 26, 1984.
The author reports that supercomputers are creeping into CAD where they can be used for structural design and simulations saving million on production costs. However, their use is limited due to their high costs of operation and general unfriendliness.

[76] Ostlund, N. S., Hibbard, P. G., and Whiteside, R. A.
A Case Study in the Application of a Tightly Coupled Multiprocessor to Scientific Computations,
Parallel Computations, Vol. 1, Computational Techniques, Pages 315-363. Academic Press, 1982.

[77] Parkinson,D.
Using the ICL DAP.
*Computer Physics Communications*26(3-4):227-232, June, 1982.
A report on the performance and characteristics of the ICL Distributed Array Processor after one year of operations.

[78] Parkinson,D. and Liddell,H.M.
The Measurement of Performance on a Highly Parallel System.
*IEEE Transactions on Computers*C-32(1):32-37, Jan, 1983.
The problems of measuring the performance of a highly parallel multiprocessor system are presented in relation to the conventional methods used for serial processors. The authors demonstrate the importance of choosing parallel algorithms which make the best use of the parallelism of the hardware by citing results obtained by DAP users while studying a wide range of application problems.

[79] Parkinson,D. and Sylwestrowicz,J.
DAP in Action.
*ICL Tech. Journal*3(3):330-344, May, 1983.
This paper gives brief summaries of the accounts of 12 different applications of the ICL Distributed Array Processor. These demonstrate the ease with which the machine can be used to attack a wide variety of types of problem and high level of computer power which the novel architecture provides.

[80] Peterson,V.L.
Application of Supercomputers to Computational Aerodynamics.
Technical Report NASA-TM-85965, NASA,
1984.
Factors pacing advances in computational aerodynamics are identified, including the amount of computational power required to take the next major step in the discipline. Additionally, NASA's Numerical Aerodynamic Simulation Program - with is 1988 target of achieving a sustained computational rate of 1 billion floating point operations per second and operating with a memory of 240 million words - is discussed in terms of its goals and its projected effect on the future of computational aerodynamics.

[81] Peterson, V. L.
Impact of Computers on Aerodynamics Research and Development.
In *Proceedings of the IEEE 72*, Pages 68-79. IEEE, 1984.
This paper describes the formulation of mathematical models of aerodynamics problems. It discusses the role of computational capability in advancing the state of the art of aeronautics and aerodynamics research.

[82] No Author Given.
Physics Today.
American Institute of Physics, 1984.
This issue contains four articles on application of very high performance computing in the modeling behavior of equilibrium fluids.

[83] No Author Given.
.
Physics Today, 1984.
This special issue of Physics Today gives a review of the general status of applications of computers in physics.

[84] Potter,J.L.
Image Processing on the Massively Parallel Processor.
*Computer*16(1):62-67, Jan, 1983.
The author demonstrates that the MPP, with an execution rate of 6 billion 8-bit adds and 1.8 billion 8-bit multiplies is particularly adept at pattern recognition and image processing. The machine's SIMD architecture can calculate 16,284 statistical pattern recognition results, extract features from a 128 by 128 subimage and search recognition and production rules in parallel.

[85] Pottle,C.
Solution of Sparse Linear Equations Arising from Power System Simulation on Vector and Parallel Processors.
*ISA Transactions*18(3):81-88, Mar, 1979.
Modern power system simulation algorithms rely on the repeated solution of large, sparse sets of linear algebraic equations. A reordering to produce a coefficient matrix in block bordered diagonal form (BBDF) is examined. Among the conclusions reached are that array processors have an immediate application in power systems simulation work, that certain directions in parallel machine organization need longer term investigation and that the vector supercomputers such as the CRAY-I do not appear to be well-suited to these problems.

[86] No Author Given.
.
In *Proceedings of the IEEE*, IEEE, 1984.
Special issue on supercomputers-their impact on science and technology. This special issue of IEEE Transactions includes a number of articles on the numerical applications of very high performance computers. Each of the articles gives an interesting discussion of a particular application area and often includes a very substantial list of technical references.

[87] Rodrigue, G., Hendrickson, C. and Pratt, M.
An Implicit Numerical Solution of the Two-Dimensional Diffusion Equation and Vectorization Experiments,
Parallel Computations, Vol. 1, Computational Techniques, Pages 101-128. Academic Press, 1982.

[88] Rose, D. J.
Numerical Analysis for VLSI Simulation: Pointers and Reflections,
Frontiers of Supercomputing. University of California Press, .
This paper discusses the modeling of the physical behavior of VLSI devices. The execution behavior of these devices are described by a coupled set of non-linear partial differential equations. The electrical behavior of an integrated circuit is captured by solving a large system of sparse non-linear algebraic differential equations. The combination of these models leads to a very complex computational problem. The work in VLSI modeling exploiting this representation is described in this article.

[89] Rubbert, P. E.
The Impact of Computational Methods on Aircraft Design.
Cray Channels:63-65, 1984.
This article discusses applications of very high performance computing in computational fluid dynamics. It gives sketches of two applications of computational fluid dynamics in the development of the Boeing 757, 767 and 737-300 aircraft. This paper is a condensation of a larger, more comprehensive article by the same author presented at the Proceedings of the A.I.A.A. Atmospheric Flight Mechanics Conference in August, 1983.

[90] Schatz, W. and Verity, J. W.
DARPA's Big Push in A.I.
*Datamation*30:48-50, 1984.
This is another brief article describing the DARPA Strategic Computing Program.

[91] Schommer, N.
Crashing Images.
Discover:74-76, October, 1984.
This article describes the work by the automobile manufacturers in trying to improve through computer simulation of automobile crashes. The effectiveness of the process is limited by computational power.

[92] Schwartz, J. T.
Mathematical Problems in Robotics,
Frontiers of Supercomputing. University of California Press, .
This article describes the selection of computational problems which must be solved in order to have general purpose, flexible, intelligent robotics. The work includes specification of very high performance sensor manipulations,

very fast execution of geometric and physical computations and effective coupling of this information to the real world via symbolic, intelligent behavior.

[93] Siebel, G.
Computational Chemistry: A Biomacro Molecule.
*Cray Channels*Spring:3-7, 1985.
This article discusses the synthesis and evaluation of molecular structures with desired physical and/or biological properties. It defines the techniques to accomplish this research and gives some of the results which have been obtained.

[94] Smarr, L. L.
An Approach to Complexity: Numerical Computations.
*Science*228:403-408, .
This article describes an application of very large scale computing and characterizes the quality and character of the information which can be obtained from a precise solution to a partial differential equation.

[95] Solem,J.C.
MECA: A Supercomputer for Monte Carlo.
Technical Report LA-10005, Los Alamos National Labs.,
1984.
The author compares discrete-ordinates and Monte-Carlo techniques for solving integrodifferential equations, and compares their relative adaptability to vector processors. The utility of multiprocessors for Monte Carlo calculations is discussed and MECA's simple architecture is described.

[96] No Author Given.
RNA Modeling for Biotechnology.
*Cray Channels*Spring:8-11, 1985.
This article describes the computational modeling of RNA molecules. This articles sketches some of the results which have been obtained with this approach to structure of molecular sequences.

[97] Stevenson,D.K.
Numerical Algorithms for Parallel Computers.
In *1980 AFIPS Conference Proceedings*, Pages 357-361. May, 1980.
The author indicates how the main features of vector and array processors affect the formulation of efficient parallel algorithms. This leads to considerations involving information flow during a parallel algorithm and its affects on bandwidth requirements. The concluding section suggests the relevance of this work to the explosive growth in integrated circuit technology.

[98] Strikwerda, J. C.
A Time-Split Difference Scheme for the Compressible Navier-Stokes Equations with Applications to Flows in Slotted Nozzles,
Parallel Computations, Vol. 1, Computational Techniques, Pages 251-267. Academic Press, 1982.

[99] Sugarman, R.
Superpower Computers,
IEEE Sprectrum, Pages 28-34. , 1980.
This article is a survey of applications of very high performance computers as it was seen in 1980.

[100] Sun, Marjorie.
The Pentagon's Ambitious Computer Plan.
*Science*222:1213-1215, .
This article is a brief survey of the DARPA plans for the development of the Strategic Computing Initiative. It focuses on the applications of artificial intelligence in battle.

[101] Swarztrauber, P. N.
Vectorizing the FFTs,
Parallel Computations, Vol.1, Computation Techniques, Pages 51-84. Academic Press, 1982.

[102] Teitelman, R.
Designs for Living.
Forbes:298-302, November, 1984.
This article describes applications of very high performance computing to the design and development of RNA and DNA molecules for particular purposes. In particular, it describes the work of the molecular genetics company and of Robert Teitelman of the N.I.H.

[103] Thapar, N.
Solving a Drug Problem.
Computing:19-20, December, 1984.
This article is a survey of work done in Britain to discover new drugs through the applications of very high performance computing.

[104] Walker,R.B., Hay,P.J. and Galbraith,H.W.
Supercomputer Requirements for Theoretical Chemistry.
Technical Report CONF-800814-26, Los Alamos National Labs., 1980.
Examples of theoretical chemistry problems now being implemented on the Cray-1 at Los Alamos are given. The problems are taken from the fields of molecular electronic structure calculations, quantum reactive scattering calculations, and quantum optics.

[105] Walker,R.B., Hay,P.J. and Galbraith,H.W.
Supercomputer Requirements for Theoretical Chemistry,
ACS Symposium Series, Pages 47-63. American Chemical Society, Washington, D.C., 1981.
Annotation - TBD.

[106] Welsh, J. G.
Geophysical Fluid Simulation on a Parallel Computer,
Parallel Computations, Vol. 1, Computational Techniques, Pages 269-277. Academic Press, 1982.

[107] Willerton,D.L.
Cray/VAX Interactive Graphics System.
Technical Report LA-UR-83-3397, Los Alamos National Lab., 1983.
A new graphics system to suport analysis of results of supercomputer codes brought into production at Los Alamos National Laboratory is presented. The system is described from three points of view: hardware, software, and software functions.

[108] Williamson, D. L. and Swarztraber, P.
A Numerical Weather Prediction Model-Computational Aspects on the Cray 1.
In *Proceedings of the IEEE 72*, Pages 56-57. 1984.
This paper presents a case study in the design and implementation of the numerical weather prediction model on a supercomputer. The model is developed and the numerical solution is developed.

[109] Wilson, K. G.
The Role of Universities in Very High Performance Computing,
Frontiers of Supercomputing. University of California Press, to be published.
This article takes the position that it is very important for the universities to have access to very high performance computing in order to develop the applications of very high performance computing. In particular, the role of the access to very high performance computers for researchers doing applications is stressed.

[110] Wilson, K. C.
Experiences with a Floating Point Systems Array Processor,
Parallel Computations, Vol. 1, Computational Techniques, Pages 279-313. Academic Press, 1982.

[111] Woodward, P. R.
Trade-Offs in Designing Explicit Hydrodynamical Schemes for Vector Computers,
Parallel Computations, Vol. 1, Computational Techniques, Pages 153-171. Academic Press, 1982.

[112] Wood,L.L.
Conceptual Basis for Defense Applications of Supercomputers.
Technical Report UCID-19746, Lawrence Livermore National Labs.,
1983.

This report concludes that the basic, long-term trends in American society suggest that the useful military power and thus the ability to influence world events of the United States will depend in the future to an ever smaller degree on Americans under arms. It is suggested that aggressive use of battlefield robotics, utilizing integrated state-of-the-art supercomputers can perform two key functions: target acquisition and weapon delivery.

3. ARCHITECTURES

[1] Alexander,T.
Computing with Light at Lightning Speed.
*Fortune Int.*110(2):74-78, July 23, 1984.
The article reports on the possibility of using optical processing, an emerging technology that uses light waves to do the work of electrons. Such processors would not only be inexpensive but would be able to recognize objects or crunch numbers faster than today's supercomputers.

[2] Alexander, G.
SSI/MSI/LSI/VLSI/USLI.
*Mosaic*15:10-17, 1984.
This article describes the development of chip architectures over the past decade and projects into the future.

[3] Babb,R.G.
Parallel Processing with Large-Grain Data Flow Techniques.
*Computer*17(7):55-61, July, 1984.
The paper presents a model of computing that represents a compromise between data flow and traditional approaches to developing large-scale applications systems for multiprocessor supercomputers. It addresses the single model of computing that can be mapped onto a variety of high-speed architectures.

[4] Baer,J.L.
Multiprocessing Systems.
*IEEE Transactions of Computers*C-25(12):1271-1277, Dec, 1976.
Surveys the state of the art in the design and evaluation of multiprocessing systems. The additional complexity of the software for control, synchronization, efficient utilization and performance monitoring of multiple processors is emphasized.

[5] Baer,J.L.
Computer Systems Architecture.
Computer Science Press, Potomac, Md., 1980.
This book is a comprehensive study of computer architecture from the introductory level to advanced research concepts. The author gives a historical survey of computer system architectures, arithmetic algorithms and a comprehensive treatment of CPU, memory, I/O and control unit design. The book describes a complete hierarchy of machines from microcomputers to supercomputers and culminates with a consideration of probable future trends in computer architecture.

[6] Baqai,I.A. and Lang,T.
Reliability Aspects of the ILLIAC-IV Computer.
In *Proc. of the 1976 Inter. Conf. on Parallel Processing*, Pages 123-131. IEEE Computer Society, New York, N.Y., Aug, 1976.
The use of dynamic redundancy techniques is proposed to make the ILLIAC-IV fault tolerant. Two different replacement schemes describing the system recovery after a PE fails are presented. In addition, the circuit implementation of the recovery mechanism are discussed.

[7] Barnes,G.H., Brown,R.M., et. al.
The ILLIAC IV Computer.
*IEEE Transactions on Computers*C-17(8):746-757, Aug, 1968.
The ILLIAC IV is a parallel-array computer containing 256 processing elements interconnected in a grid pattern. Special features include multiarray processing, multiprecision arithmetic and fast data-routing interconnections. This is the definitive paper on the ILLIAC IV computer.

[8] Batcher,K.
MPP - A Massively Parallel Processor.
In *Proc. of the 1979 Inter. Conf. on Parallel Processing*, Pages 249. IEEE Computer Soc. Press, New York, N.Y., Aug, 1979.
Annotation - TBD.

[9] Bell, G. C.
Multis: A New Class of Multiprocessor Computers.
*Science*228:462-466, 1985.
This article describes a new class of computers based on multiple processors. It points to the many applications which can be speeded up by parallel computation and the tremendous cost advantages that one expects to get from these systems.

[10] Broughton,J.M., Farmwald,P.M. and McWilliams,T.M.
S-1 Multiprocessor System.
Technical Report UCRL-87494, Lawrence Livermore National Lab., 1982.
This paper describes the S-1 multiprocessor system. The hardware organization, its implications, and software supporting the efficient utilization of the multiprocessor are discussed.

[11] Browne, J. C.
Parallel Architectures for Computer Systems.
Physics Today 37:28-37, 1984.
This article describes the history and the current concepts for development of new generations of very high performance computers.

[12] Bucher,I.Y. and Moore,J.W.
Comparative Performance Evaluation of Two Supercomputers: CDC Cyber-205 and CRI Cray-1.
Technical Report LA-UR-81-1977, Las Alamos National Lab., 1981.
This report compares the preformance of CDC's newest supercomputer, the Cyber-205, with the CRI Cray-1. Results are presented of tests to determine the speed of basic arithmetic operation, of runs using our standard benchmark programs, and of runs using three codes that have been optimized for both machines. It is concluded that the speed of the Cyber-205 for memory-to-memory operations on vectors stored in consecutive locations is considerably faster than that of the Cray-1.

[13] Budnik,P. and Kuck,D.J.
The Organization and Use of Parallel Memories.
*IEEE Transactions of Computers*C-20(12):1566-1569, Dec, 1971.
Some of the important questions concerning the design and use of parallel memories in computer systems with high execution rates are discussed,

[14] Davis,R.L.
The ILLIAC IV Processing Element.
*IEEE Transaction on Computers*C-18(9):800-816, Sept, 1969.
The processing element of the ILLIAC IV is described and analyzed.

[15] Davis,A.L.
Computer Architecture.
*Spectrum*20(11):94-99, Nov, 1983.
The author discusses the main issues involved in designing and building concurrent computer architectures. Three prototype machines each using a different control scheme are described and analyzed.

[16] Dennis, J. B.
Data Flow Ideas in Future Supercomputers,
Frontiers of Supercomputing. University of California Press, .
This article reports the concepts developed for implicit synchronization and their potential impact on future very large scale computations.

[17] Dennis,J.B.
Data Flow Supercomputers.
*Computer*13(11):48-56, Nov, 1980.
The data flow concept and two contrasting data flow architectures, the data flow multiprocessor and the cell block architecture, are presented and described. The different approaches taken by the two architectures in communicating information between system components are analyzed.

[18] Drogin,E.M.
Pipeline Architecture Battles Array.
*MSN Microwave System News*10(10):92-96, Oct, 1980.
Discusses pipeline and array processing concepts, emphasizing practical distributed processors.

[19] Eddleson, E.
Computers at the Forge.
This is a non-technical article describing the elements in a total system architecture, including chip technologies, networking in A.I., which need to be improved to allow the development of new generations of computer systems.

[20] Efe,K.
Task Allocation in Supersystems.
Technical Report LU/DCS/R-173, Leeds University, 1983.
The report considers the implications of interconnecting many (possibly several hundreds) of mini- or micro-computers to obtain enormous computational power. The report discusses the program allocation scheduling problems with various types of systems and suggests various algorithms. Future trends in the design of supersystems are discussed as well as some outstanding research problems.

[21] Farmwald,P.M., Bryson,W. and Manferdelli,J.L.
Signal Processing Aspects of the S-1 Multiprocessor Project.
Technical Report CONF-800719-7, Lawrence Livermore National Labs., 1980.
This paper gives a general description of the S-1 Mark IIA uniprocessor and somewhat more extensive discussion of its signal processing hardware and hardware-executed algorithms. The focus is a discussion of the felicitous properties of the uniprocessor components that make the S-1 system a powerful general-purpose digital signal processor.

[22] Gottlieb, A.
The New York University Ultracomputer,
Frontiers of Supercomputing. University of California Press, .
This article describes the architecture and some applications which have been made of the New York University designed ultra-computer.

[23] Hanuliak,I.
Computer Systems from the Viewpoint of Architecture.
*Automatizace*23(8):205-209, Aug, 1980.
The main features of uniprocessors, array processors, multiprocessors, associative processors and multicomputers are outlined and illustrated with typical commercial systems. Future development trends until 1990

are presented in a graphical form.

[24] Hiatt, B. and Gwinne, P.
Parceling the Power.
*Mosaic*15:36-43, 1984.
This article describes the research problem formulations which will be needed to make parallel computers effectively useable.

[25] Hogen,D.J.
Computer Performance Prediction of a Data Flow Architecture.
Master's thesis, Naval Postgraduate School, 1981.
This thesis uses a Petri net-base methodology to evaluate the performance attainable by supercomputers based on data flow architecture. Results show that for such architectures to achieve their goal of high speed computation, intelligent multiprogramming schemes need to be developed. One such scheme, based on the notion of a 'concurrency vector' is introduced.

[26] Hoshind,T., Kageyama,T., Shirakawa,T., et. al.
Highly Parallel Processor Array 'PAX' for Wide Scientific Applications.
In *Proc. of the 1983 Inter. Conf. on Parallel Processing*, Pages 95-105. IEEE Computer Soc. Press, Silver Springs, Md., Aug, 1983.
The architecture of PAX is characterized by end-around nearest-neighbor interprocessor connection, data broadcasting capability, asynchronous MIMD operation and global synchronization control. The authors give execution times for various scientific processes both for the current 32 processor machine as well as for a proposed super parallel machine with over 1000 processors.

[27] Hoshind,T., Kageyama,T., Shirakawa,T., et. al.
PACS: A Parallel Microprocessor Array for Scientific Calculations.
*ACM Transactions on Computer Systems*1(3):195-221, Aug, 1983.
PACS-32, a prototype supercomputer for scientific calculations is described. The machine provides a synchronizable MIMD processor array, efficient nearest neighbor communication and a high-level language.

[28] No Author Given.
ICOT,Outline of Research and Development Plans for Fifth Generation Computer Systems. Institute for New Generation Computer Technology.
This report emphasizes themes of Japan's R&D efforts in the area of fifth-generation computers. It also looks at the system's functions and overall plans.

[29] Jensen,C.
Taking Another Approach to Supercomputing.
*Datamation*24(2):156-172, Feb, 1978.
The Burroughs' large-scale scientific processor is described.

[30] Jordan, T. L.
A Guide to Parallel Computation and Some Cray-1 Experiences,
Parallel Computations, Vol. 1, Computational Techniques, Pages 1-49. Academic Press, 1982.

[31] Kaneda,Y.
Parallel Processors.
*Systems and Control*27(8):508-516, Aug, 1983.
The speeding-up of parallel processing by pipeline control with advance order control, memory interleaving and calculation units is described. As examples of pipeline control computers the STAR-100, CYBER-205, CRAY-I and AP-120B System are treated in detail and compared. Vectorization techniques in application programs are discussed.

[32] Karplus,W.J.
Parallelism and Pipelining in High-Speed Digital Simulators.
In *10th IMACS World Cong. on Sys. Siml. and Sci. Comp.*, Pages 272-274. IMACS, New Brunswick, N.J., Aug, 1982.
The author considers the progress of supercomputers, peripheral array processors and microcomputer networks towards attaining the high execution rates required by scientific and engineering simulations.

[33] Kartashev,S.I. and Kartashev,S.P.
Problems of Designing Supersystems with Dynamic Architectures.
*IEEE Transactions on Computers*C-29(12):1114-1132, Dec, 1980.
The paper discusses the basic concepts of designing a supercomputer with dynamic architecture. It is shown how such a system can be constructed from Dynamic Computer Groups. The authors discuss hardware components, design techniques and reconfiguration strategies.

[34] Kawanobe, K.
Present Status of the Fifth Generation Computer Systems Project.
*ICOT Journal*5:13-21, October, 1984.
This article provides an overview of the entire Fifth Generation Computer Systems Project, the goals of individual research themes and the current research status as viewed from a macro standpoint.

[35] Kozdrowicki,E.W. and Theis,D.J.
Second Generation of Vector Supercomputers.
*Computer*13(11):71-83, Nov, 1980.
The improvements in second-generation supercomputers over the first are stated. In addition, three second-generation supercomputers, the Burroughs Scientific Processor, the Cyber 205 and the Cray I, are described, analyzed and compared in detail.

[36] Kozdrowicki,E.W.
Supercomputers for the Eighties.
*Digital Design*13(5):94-103, May, 1983.
The author looks at trends in the architecture of array or vector processors.

[37] Kuck, D. J., Laurie, D., Cytron, R., Same, H. A., and Gajski, D.
Cedar Project,
Frontiers of Supercomputing. University of California Press, .
This article describes the concept basis of the Cedar supercomputer which is being designed and implemented at the University of Illinois. The Cedar is unique in having a combination of hardware and software concepts as its basis.

[38] Kuck,D.J. and Padua,D.A.
High-Speed Multiprocessors and Their Compilers.
In *Proc. of the 1979 Inter. Conf. on Parallel Processing*, Pages 5-16. IEEE Computer Soc. Press, New York, N.Y., Aug, 1979.
Annotation - TBD.

[39] Kung, H. T.
Special Purpose Supercomputers.
Technical Report , Carnegie-Mellon,
1984.
This paper describes the application of systolic arrays in building very high performance machines.

[40] Labrecque, M.
The Many Ways Data Must Flow.
*Mosaic*15:18-27, 1984.
This is an informal article describing the parallel computer architectures which are arising in research projects.

[41] Levine,R.D.
Supercomputers.
*Scientific American*246(1):118-135, Jan, 1982.
The article describes in detail the Cray-I and the Cyber 205 and gives a brief description of the ILLIAC-IV, the S-1 and the HEP. The use of pipelining and multiprocessing by these machines is brought out in the article.

[42] Manuel, T.
Advanced Parallel Architectures Get Attention as Way to Faster Computing.
*Electronics*56(12):105-106, June 16, 1983.
This article discusses very large-scale integration and the effect of funding on advances in parallel designs.

[43] Manuel, T.
Cautiously Optimistic Tone Set for 5th Generation.
*Electronics Week*December3":57-63, 1984.
This article reviews progress made over the past three years in Japan's fifth-generation computer project. Goals of the program are discussed. Also included is information on parallel-processing architectures.

[44] Martin,H.G.
A Discourse of a New Super Computer, PEPE.
In *Conf. on High Speed Computer and Algorithm Organization*, Pages 101-112. Academic, London, England, Apr, 1977.
PEPE, a parallel array processor with distributed and associative logic, is capable of floating point computation rates in excess of 500 MIPS. The distributed and associative logic was modeled after the Bell Lab integrated circuit PEPE. The parallel logic has been augmented with three sequential processors each with separate data and program memories.

[45] No Author Given.
And Now, an 'Affordable Supercomputer'.
Business Week Industrial Ed.:164, Nov, 1984.
Convex Computer's Convex Model C-1 is introduced. It is shown to run 3-5 times as fast as Digital Equipment's new supermini, the VAX 8600, and the same as DEC's lowest-priced version.

[46] No Author Given.
A Plethora of Projects in the U.S. Try Data-flow and other Architectures.
Electronics:107-110, June 16, 1983.
Research in advanced parallel architectures being conducted in U.S. universities is reviewed. Efforts now underway at the University of Illinois, Purdue, the University of North Carolina Chapel Hill are discussed.

[47] No Author Given.
SuperCPUs: Japanese Close In.
Electronic Engineering Times:1,12, Mar 16, 1984.
The article reports that Fujitsu's VP-2000 has been impressive in benchmark tests at NASA's Ames Research Center. One strong point indicated by the benchmarks is the vectorizing power of the VP-2000's compiler; it can vectorize significantly more DO loops than the CRAY-XMP (although the Version 1.11 NASA/Ames CRAY FORTRAN compiler, CFT, allows only one of the two CRAY CPUs to be used).

[48] Parkinson,D.
The Distributed Array Processor (DAP).
*Computer Physics Communications*28(4):325-336, Feb, 1983.
The Distributed Array Processor, a system consisting of 4096 processors

working in unison under the control of a single master processor is described.

[49] Requa,J.E. and McGraw,J.R.
The Piecewise Data Flow Architecture: Architectural Concepts.
*IEEE Transactions on Computers*C-32(5):425-438, May, 1983.
This paper presents the design and a brief analysis of the Piecewise Data Flow Computer (PDF). The PDF is a heterogeneous multiprocessor system having both SIMD and MIMD characteristics. Concurrency can be exploited in three different ways: simultaneous execution of independent basic blocks, simultaneous execution of independent instructions within a basic block and intrinsic array operations.

[50] Rosenberg, R.
Supercube.
*Electronics Week*58(6), February 11, 1985.
This article describes the Intel personal supercomputer based on the Caltech Cosmic Cube

[51] Rudy,T.E.
Magaflops from Multiprocessors.
In *Proc. of the 2nd Rocky Mountain Symp. on Microcomputers*, Pages 99-107. IEEE Computer Soc. Press, New York, N.Y., Aug, 1978.
The author points out that the cost and performance level of vector or array processor machines is dependent on resolving such problems as resource contention, task partitioning and synchronization. Before these issues can be resolved a high-level language specification is probably unavoidable.

[52] Sahni,S.
Scheduling Supercomputers.
Technical Report TR-83-3, Minnesota University, 1983.
Good heuristics to schedule tasks on supercomputers are developed. Supercomputers comprised of multiple pipelines as well as those comprised of asynchronous multiple processors are considered.

[53] Schindler,M.
Multiprocessing Systems Embrace Both New and Conventional Architectures.
*Electronic Design*32(6):97-130, Mar, 1984.
The author points out that multiprocessing architectures are now so diverse that finding the best one for a given application is a major chore. Serving as a guide, this report on multiprocessing points out safe areas, as well as pitfalls. It explores several popular schemes for classifying multiprocessors: the degree of coupling, how data and instructions flow through the system and how processors exchange information.

[54] Seitz, C., and Matisoo, J.
Engineering Limits on Computer Performance.
Physics Today 37:38-52, 1984.
This article describes the fundamental limitations on device technology which arise from the fundamental principles of physics. It also projects what performance can be expected in the near future from VLSI technology.

[55] Simmons,G.H.
TDMP: A Data Flow Processor.
PhD thesis, Michigan State University, 1981.
This research investigates a multiple processor computer structure that has a higher bandwidth than comparable single-processor machines yet is more flexible than existing fixed-array multiprocessors. The basic approach taken here was to develop, model, simulate and then analyze a computer structure called TDMP - an acronym for Time Division Multiple Processor.

[56] Smith,M.G., Notz,W.A. and Schischa,E.
The Question of Systems Implementation with Large-Scale Integration.
*IEEE Transactions on Computers*C-18(8):690-694, Aug, 1969.
Some of the questions asked about LSI technology in the late 60s are discussed. These include how LSI will affect hardware costs, system costs and system design.

[57] Smith,B.J. and Fink,D.J.
Architecture and Applications of the HEP Multiprocessor Computer System.
In *Proc. of the 1982 Conf. on Peripheral Array Processors*, Pages 159-170. SCS, La Jolla, Ca., Oct, 1982.
The HEP computer system, a large-scale scientific parallel computer employing shared resource MIMD architecture is described. The hardware and software facilities provided by the system are presented and techniques found to be useful in programming the system are discussed.

[58] Snyder,L.
Supercomputers and VLSI: The Effect of Large Scale Integration on Computer Architecture.
Technical Report TR-84-08-05, Washington University at Seattle, 1984.
The use of VLSI technology to build supercomputers is analyzed in depth. The benefits of VLSI are reviewed and the liabilities are explored.

[59] Su,S.P.
Pipelining and Dataflow Techniques for Designing Supercomputers.
PhD thesis, Purdue University, 1982.
This thesis investigates some pipelining and dataflow techniques for designing supercomputers. In the pipelining area, new techniques are developed for

scheduling vector instructions in a multi-pipeline supercomputer and for constructing VSLI matrix arithmetic pipelines for large-scale matrix computations. In the dataflow area, a new approach is proposed to dispatch high-level functions for dependence-driven computations.

[60] Swartzlander,E.E. and Gilbert,B.K.
Supersystems: Technology and Architecture.
*IEEE Transactions on Computers*C-31(5):399-409, May, 1982.
The paper examines three of the classical design options for supercomputers: high-speed monoprocessors, array processors and distributed processors. The paper presents metrics applicable to interconnection network performance, cost and overall quality estimation. The paper concludes with a case study of the Dynamic Spatial Reconstructor, a processor developed for ultra-high-speed tomography.

[61] Thurber,K.J.
High-Performance Parallel Processors.
In *Proc. of the Soc. of Photo-Optical Instr. Engineers*, Pages 45-59. Soc. Photo-Optical Instr. Engineers, Bellingham, Wa., Aug, 1978.
This paper presents a survey and tutorial of high-performance parallel processors. It covers the subjects of multiprocessors, pipeline processors, array/associative processors and scientific processors. Architecture summaries are provided.

[62] Treleaven,P.C.
Future Computers: Logic, Data Flow, ..., Computer Flow?
*Computer*17(3):47-55, Mar, 1984.
Likely candidates for future architectures are listed; they are: fifth-generation computers, supercomputers, VSLI processor architectures and integrated communications and computers. This paper examines each of these potential candidates for the future computer and attempts to determine the 'ideal' choice or combination of choices that will provide both computational control and system flexibility.

[63] Vick,C.R., Kartashev,S.P. and Kartashev,S.I.
Adaptable Architectures for Supercomputers.
*Computer*13(11):17-35, Nov, 1980.
The advantages and the shortcomings of dynamic architectures are discussed. In particular, the properties and possible improvements to pipeline, array processors and multicomputer/multiprocessor architectures are described. The authors show how such architectures, with many of the improvements mentioned, can be constructed from Dynamic Computer Groups.

[64] Waldrop, M. M.
Artificial Intelligence in Parallel.
Science 25:608-610, 1984.

This article describes the computer architectures which are being developed for parallel structuring and does very high performance computation for A.I. applications.

[65] Widdoes,L.C.
S-1 Project: Developing High-Performance Digital Computers.
Technical Report CONF-800201-5, Lawrence Livermore National Labs., 1980.

The author reports that a multiprocessor, the S-1, with at least ten times the computational power of the Cray-1 is being designed and implemented. The first step is to develop a general-purpose uniprocessor with a performance level comparable to the Cray-1; the multiprocessor will then be made up of 16 of these uniprocessors, sharing a main memory.

4. BUSINESS AND MARKETING ASPECTS

[1] Alexander,J.
Reinventing the Computer (Parallel Processing).
*Fortune*109(5):62-70, Mar 5, 1984.
The author notes that today's supercomputers are too slow for many potential tasks; some tasks taking weeks or months. Currently, companies and nations are scrambling to exploit a faster, smarter technology known as parallel processing in hopes of dominating what in a decade should become a $500 billion a year industry. This intense research activity should produce a new generation of high-speed computers based on parallel computing which will open up many new application areas for computers.

[2] Alexander, T.
Cray's Way of Staying Super-duper.
*Fortune*111(6):66-68, 72,76, March 18, 1985.
This article examines Cray Research's strategy to stay ahead of its competitors in the supercomputer field.

[3] Benoit, E.
Filling the Gap.
Forbes:166-170, March, 1985.
This article describes the rise of new systems which are being thought to have a very high relative performance and to have the performance of a supercomputer of one generation back for a tenth of the price of a current supercomputer. The focus is on the potential and on the marketplace.

[4] Buzbee, B. L., Netropolis, N. and Sharp, D. H.
Frontiers of Supercomputing.
In *Frontiers of Supercomputing*, Pages 64-80. Los Alamos, 1983.
This article reports the principle conclusions reached at a meeting sponsored by Los Alamos and NSA on the future of very large scale computation.

[5] Buzbee, B. L. and Sharp, D. H.
Perspective on Supercomputing.
*Science*227:591-597, 1985.
This article gives a brief look at the current status of supercomputers and application of supercomputers in the United States. It discusses the characteristics of a large modern supercomputer facility. It defines some of the changes in the architecture of future generations of supercomputers and discusses some applications from Los Alamos Scientific Laboratories.

[6] Carlyle, R. E.
Here Come the Crayettes.
*Datamation*31:40-42, 1985.
This article describes the surprise of the marketplace and the companies which

are making very high performance computers at a moderate. That is, computers which have a very high price/performance ratio and are close to the top end of the line.

[7] Cook, J.
War Games.
*Forbes*132(6):108,110, September 12, 1983.
This article describes Cray Research's market strategies to maintain its position in supercomputer arena. Products including the Cray-1, 1-M and the X-MP are reviewed.

[8] Corrigan, R.
Inman's Innovation.
National Journal:513, March 5, 1983.
This article describes one high-tech response by corporate America (MCC) to the Japanese technological challenge.

[9] Feigenbaum, E.A. and McCorduck, P.
Japan's MITT Strike.
*Savvy*4(5):38,40,42,44, May, 1983.
This article discusses how, according to the Ministry of International Trade and Technology, the Fifth Generation computer systems will secure both Japan's economic future and its worldwide technological leadership.

[10] Fernbach, S.
The IEEE Supercomputer Committee Report.
Technical Report , IEEE,
1983.
This report is a product of a committee appointed by the United States Activities Board of the IEEE. It gives a summary of importance of supercomputers, both in economic and national security terms. It summarizes current applications and the current technology for supercomputers.

[11] Gerola, H. and Gomory, R. E.
Computers in Science and Technology: Early Indications.
Science 25:11-18, 1984.
This article describes the different modes in which scientists and engineers have used large scale computing. They project an overall twenty percent growth annually in the capability of computer systems. They state that this technology has 'already had such major impacts on scientific research engineering that it is important to try to predict its evolution and uses.'

[12] Johnson, J.
Cray and CDC Meet the Japanese.
*Datamation*30:32-42, 1984.
This article describes the relative status in 1984 of CDC, Cray, Fujitsu, NEC

and Hitachi vector supercomputer systems. It reported some of the results from Mendez's Livermore benchmarks.

[13] Kartashev,S.P. and Kartashev,S.I.
Supersystems for the 80's.
*Computer*13(11):11-14, Nov, 1980.
An introduction to the Novmber, 1980 issue of Computer dedicated to supercomputers. The authors note that recent advances in LSI technology have made adaptable and data flow architectures possible. Such architectures support distributed processing and greatly increase execution rates.

[14] Kurita, S.
Supercomputers Battle Japan Inc. on its Own Turf.
*Electronic Business*10(19):70,73, December 10, 1984.
This article looks at Cray Research's supercomputer market. Also market strategies being implemented by the Japanese supercomputer companies.

[15] Lineback, J.R.
Japan Looms as a Key Market for U.S.-made Supercomputers.
*Electronics Week*57(24):18-19, September24", 1984.
This article describes how Japan is becoming a potential market for U.S. supercomputers and how the various U.S. supercomputer companies are evaluating this possibility.

[16] Manuel, T. and Cohen, C. L.
Computers of All Categories Increase in Performance: Fifth-Generation Machines and Supercomputers Hottest.
*Electronics Week*57(24):58-60, September 24, 1984.
The developers of supercomputers and next-generation computers are concentrating on large increases in computational power, information-processing, power and storage power over what the current machines offer. As Japanese companies prepare their models for market, the U.S. supercomputer leaders are developing their next products.

[17] Mendez, R. and Orszag, S.
The Japanese Supercomputer Challenge.
*Datamation*30(7):113,114,116,119, May 15, 1984.
This article offers a comparison of Japanese and U.S. supercomputers, the technology itself and its use. It also reviews current state of the U.S. market and competitors involved.

[18] No Author Given.
Supercomputers are Breaking Out of a Once Tiny Market.
Business Week Industrial Ed.:164, Nov, 1984.
It is reported that supercomputer sales will grow from $300 million per year in 1984 to $1.5 billion per year in 1989. An accelerating shift into commercial

markets has been aided by the arrival of a bevy of new companies marketing 'near-supercomputers' and the entry of Japanese manufactures.

[19] No Author Given.
Amdahl to Introduce Two Supercomputers Made by Fujitsu Ltd.
The Wall Street Journal.
The article describes the Amdahl 1100 and 1200 computers. Both machines will be manufactured by Fujitsu of Japan and both will compatable with IBM mainframes.

[20] No Author Given.
Avionics: Smaller Firms Dominate Supercomputer Field in US.
Aviation Week & Space Technology:154-157, May 28, 1984.
The US firms, Cray Research, ETA Systems (an offshoot of CDC) and Denelcor, are compared and contrasted.

[21] No Author Given.
The Battle of the Supercomputers: Japan's All-out Challenge to the U.S.
Business Week:156,159,162,166, .
This article examines the challenge which Japan is mounting against the U.S. in supercomputers and America's response to that challenge.

[22] No Author Given.
A Fifth Generation: Computers that Think.
Business Week:94-96, December 14, 1981.
The Fifth-Generation Computer Project aims to develop computers very different from the ones now available. U.S. computer makers are keeping a close eye on this Japanese project.

[23] No Author Given.
Industry's Use of Computers Shows Big Increase.
Oil & Gas Journal:41-44, April 9, 1984.
The article reports on the current use of supercomputers in the oil and gas industry.

[24] Tyler, M.
Amdahl's Super CPU Gamble.
*Datamation*30(18):36,38,40,42, November 1, 1984.
This article reviews Amdahl's attempt to compete in the U.S. supercomputer market. It offers descriptions of the Cray and Control Data supercomputers in terms of capabilities and specific features of each machine.

[25] Uttal, Bro.
Here Comes Computer Inc.
*Fortune*107(7):82-90, October4", 1982.
This article reviews the project backed by the Japanese government to build

fifth-generation computers as Japan makes bid for leadership in advanced computer technology. It examines software problems being encountered in the project.

[26] Wilson,K.G.
Science, Industry and the New Japanese Challenge.
*Proc. of the IEEE*72(1):6-18, Jan, 1984.
The author reports that the Japanese Super-Speed Computer Project, has generated pressure for a response in the US and other countries. He identifies a number of barriers which are slowing the US effort; they are: software problems, lack of university participation in the computerization effort, poor matching of resources to needs, etc. Several recommendations are proposed, among which is the the establishment of a National Computer Communication Network for basic and applied research.

5. GENERAL ARTICLES AND SURVEYS

[1] Buzbee,B.L., Ewald,R.H. and Worlton,W.J.
Japanese Supercomputer Technology.
*Science*218(1189-1193):1189-1193, Dec, 1982.
A nontechnical article reviewing the Japanese National Superspeed Computer Project and Fifth-Generation Computer Project. The economic and security implications of these projects for the United States are discussed.

[2] Cohen, C. L.
JapaneseFifth-GenerationProgramClickingAlongRightonSchedule".
*Electronics Week*57(15):33-34, July 23, 1984.
The article reviews the progress Japan has made in the development of knowledge systems for the fifth-generation computer. Mention is also made of architecture research and application programs.

[3] Davis, D. B.
Supercomputers: A Strategic Imperative?
*High Technology*4(5):44,46,52, May, 1984.
This article examines the U.S. response to Japanese competition in supercomputers.; Included is a discussion of specific measures ie. increased funding and research and development, implemented by our government to meet this challenge.

[4] Davidson,H.L.
Some Predictions on the Performance of Future Supercomputers for Simulation and Control.
Technical Report UCRL-89969, Lawrence Livermore National Lab, 1983.
This report discusses what is meant by a supercomputer, some of the reasons why their performance is not as high as we would like, some techniques for improving this performance, predictions of how fast they will become in the near future, and what sorts of interesting and entertaining things we can do with these faster machines.

[5] Dickson, D.
Britain Rises to Japan's Computer Challenge.
*Science*220(4599):799-800, May 20, 1983.
This article describes the effort by the British government to subsidize research in information technology as part of the country's response to the Japanese fifth-generation computer program.

[6] Hendrickson,C.P.
Thinking Big.
Technical Report UCRL-90392, Lawrence Livermore National Labs., 1984.

The author reports that supercomputers have long been used at the nuclear weapons laboratories as very cost-effective computational alternatives to large numbers of nuclear tests. This paper discusses the general characteristics of these supercomputers, gives an example of the environment in which these computers are placed, and indicates what future supercomputers will look like.

[7] Iversen,W.R.
Supercomputers Find New Jobs.
*Electronics*55:75-76, July 28, 1982.
Annotation - TBD.

[8] Kahn,R.E.
A New Generation in Computing.
*Spectrum*20(11):36-41, Nov, 1983.
A nontechnical article arguing that the next generation of computers will be a mix of advanced microelectronics, artificial intelligence and multiprocessor parallel architectures. An outline of the five generations of computer and communication technologies and a glossary of computer terms is included.

[9] Knight,J.C.
Current Status of Supercomputers.
*Computers & Structures*10(1-2):401-409, Jan, 1979.
Annotation - TBD.

[10] Lemmons, P.
Japan and the Fifth Generation.
Byte:394-401, November, 1983.
Japan's efforts to develop artificial intelligence are intended to make computers user-friendly. An outline of research and development plans is included in this discussion.

[11] Malik, R.
Japan's Fifth Generation Computer Project.
*Futures*15(3):205-210, June, 1983.
This article discusses Japan's 5th Generation project and examines motives and goals of the program. One specific goal discussed is the development of the system's ability to process knowledge in a user-friendly manner.

[12] Manuel,T.
Hollywood Signs on Supercomputer.
*Electronics*55:56, Aug 11, 1982.
Annotation - TBD.

[13] Marback, W. M. and Cook,
The Race to Build a Supercomputer.
Newsweek, 1983.

This article is a review of the several programs, both government and university, that are producing new generations of supercomputers. It gives projections of the status of this research in the United States as compared to Germany.

[14] Marcom, J. and Browning, E. S.
Japan's Supercomputers: Breakthrough or Boast?
*The Wall Street Journal*LXXIV(27):34, August 8, 1984.
This article discusses the reasons why Japanese supercomputers may or may not ultimately prove to be superior to those produced here in the U.S.

[15] Marsh, P.
The Race for the Thinking Machine.
New Scientist:85-87, July 8, 1982.
This article discusses Britain's response to the Japanese fifth-generation computer project. It examines some of the difficulties the U.K. faces in this effort and the possible consequences of poor decisions on the part of the British.

[16] No Author Given.
The Race is Still Open.
*The Economist*293(7368):88,93, November 17, 1984.
In response to Japan's fifth generation computer project, numerous government-sponsored computer research projects are appearing in Europe. One specific project in Britain known as Alvey is discussed.

[17] .
Western Europe Looks to Parallel Processing for Future Computers.
Electronics:111-113, June 16, 1983.
Parallel processing projects in West Germany, France and Britain are testing advanced multiprocessor architectures. This article reviews these particular projects.

[18] No Author Given.
The Race for Computer Supremacy: Who's Ahead.
The New York Times.
The article outlines the directions being taken by the US and Japanese governments, the EEC and US manufactures in the race to build the next generation of supercomputers. The article notes that the winner may very well dominate the world economically and militarily for years to come.

[19] No Author Given.
Super Problems for Supercomputers.
Science News:200-203, Sept 29, 1984.
The article discusses the use of supercomputers to simulate various processes in molecular chemistry and biology and aid in the design of airplanes.

[20] No Author Given.
NAS Enters Vector Processing Arena with Supercomputer.
Computerworld:8, July 9, 1984.
National Advanced Systems' AS/9100 is described. The computer is seen as a supercomputer bridging the gap between general-purpose scalar mainframes and full-fledged number crunchers. Applications include aerospace development, automotive manufacturing and semiconductor design.

[21] No Author Given.
The Market for Advanced-Technology Computers.
Computer Age:1,5, July 9, 1984.
Six types of advanced-technology computers are identified and described. They are: supercomputers, near-supercomputers, array processors, associative processors, specialized processors and knowledge processors.

[22] No Author Given.
The Coming Generation of Supercomputers.
*Science Digest*91(3):74-75, Mar, 1983.
Annotation - TBD.

[23] No Author Given.
Supercomputer Capacity in the Netherlands, An Inventory of Necessity and Availability.
*Informatie*26(4):294-301, Apr, 1984.
The main manufactures of supercomputers are identified and their products described. Furthermore, the main application fields of supercomputers are listed; they are: the military, surveying, oil exploration, aircraft construction, CAD/CAM and scientific research.

[24] No Author Given.
Supercomputers Come Out Into the World.
The Economist:77-80, August11", 1984.
This article examines how supercomputers could be valuable for many industrial uses once a drop in price and a substantial increase in speed occur. The market for supercomputers could increase 30-fold by the end of the decade.

[25] Norman, C.
House Votes Florida State a Supercomputer.
*Science*224:1075-1076, June8", 1984.
In May of 1984 the House os Representatives passed a budget bill for the Department of Energy which included a $7 million appropriation for a supercomputer center at Florida State University. This article provides an overview of the Florida State proposal.

[26] Oberdorfer, D.
Japan's New Challenges.
The Washington Post:22, July 23, 1984.
The article summarizes concerns in America and other countries regarding Japan's work on supercomputers. Responses to this Japanese challenge are mentioned.

[27] Pehrs,J.
Rank and File Leap in Performance (Supercomputers Are Coming).
*Chip*14(12):16-18, Dec, 1983.
A comprehensive survey of research activities across the world is presented. The emphasis is on the dataflow concept of multiprocessor systems working in parallel. Efforts in the US, Japan and Europe and some forerunners of this new generation of computers are described. The need for higher speeds is shown on a number of scientific applications, which require hundreds and even thousands of computing hours with current technology.

[28] Sanger, D. E.
The surge in Supercomputers.
*The New York Times*134:D1,D16, March 1, 1985.
Supercomputers are in high demand and this article examines why. It also explores the growth in new supercomputer projects. Not all supercomputer market entrants have survived this frenzy of activity and this report tells why.

[29] Schefter, J.
Fifth-generation Computers.
Popular Science:79-81, April, 1983.
Japan is building a computer so advanced that its development is ranked as a national project. Hardware that can see, listen, talk, learn and make judgments is the goal for the 1990s.

[30] Seligman, M.
The Fifth Generation.
PC World:282,284,286, October, 1983.
This article examines Japan's challenge to U.S. supremacy in the computer industry and how we (the U.S.) must meet that challenge.

[31] Shimoda, H.
Fifth Generation Computer: From Dream to Reality.
*Electronic Business*10(17), November 1, 1984.
This article recounts an interview with Kozuhiro Fuchi, research director of the Institute for New Generation Computer Technology (ICOT). It reviews progress on a hardware module and a variety of current research topics.

[32] Slotnick,D.L.
Centrally-Controlled Parallel Processors.
In *Proc. of the 1981 Inter. Conf. on Parallel Processing*, Pages 16-24. IEEE Computer Soc. Press, Silver Springs, Md., Aug, 1981.
The circumstances surrounding the conception and development of centrally-controlled array processors are described (the years covered are 1953 to 1975). Some reflections on past and possible future interplay between university and government laboratories on the one hand and industry on the other are made at the conclusion.

[33] Smith, K.
U.K. Pursues Fifth-generation Computer.
*Electronics*56(11):101-102, May 31, 1983.
The British government has approved a cooperative research program aimed at developing a fifth-generation computer. This article discusses this effort.

[34] Solomon,S.
Superbrain: The Race to Create the World's Fastest Computer.
*Science Digest*91(9):42-49, Sept, 1983.
Annotation - TBD.

[35] Torrero, E. A.
Tomorrow's Computers--The Quest (Special Section).
IEEE Spectrum, November, 1983.
This issue of IEEE Spectrum reviews numerous topics relating to advanced computing. Research projects in Japan, the U.S., Great Britain and Europe are discussed. Coverage on artificial intelligence, software engineering, computer architectures and VLSI is also featured. Computer trends and next generation impacts round out this special report.

[36] Walton, P. and Tate, P.
Soviets Aim for 5th Generation.
*Datamation*30(10):53, 56,61,64, July 1, 1984.
The article summarizes Soviet efforts to compete in the race of advanced computer systems. Goals, specific projects and problems being encountered are reviewed.

[37] Wallich,P.
Designing the Next Generation.
*Spectrum*20(11):73-77, Nov, 1983.
A nontechnical article discussing the advancements necessary to realize the next generation of computers. The article points out many of the problems facing today's designers in trying to achieve those advances.

[38] Walter,G.
Intelligent Supercomputers: The Japanese Computer Sputnik.
*Journal of Information Image Management*16(10):18-22, Nov, 1983.

This article reports on the effect that Japan's Fifth-Generation Computer Project is having on the American computer and information systems industry. The American firms are intensifying their research on the production of intelligent supercomputers, a combination of computer architecture and artificial intelligence software programs. While the present generation of computers is built for the processing of numbers, the new supercomputers will be designed specifically for the solution of symbolic problems and the use of artificial intelligence software.

[39] Wilson, K. G.
Supercomputer Powerful Tool.
Austin American Statesman:J4, April 21, 1985.
This article offers a definition of supercomputers as well as possible applications of these machines. Mention is made of the Federal government's announcement regarding the establishment of supercomputer centers at four U.S. universities and what impact this plan might have on the market for computers.

[40] Worlton, J.
Understanding Supercomputer Benchmarks.
*Datamation*30(14):120,122,124,128,130, September 1, 1984.
According to this article, recent tests of supercomputers may have been unfai to American machines. An approach on how to evaluate the data to get a clearer picture of what these tests really mean is offered.

[41] Yasaki, E.K.
Japan Goes for the Gusto.
*Datamation*27:40,44,47,50,55, August, 1981.
This article reviews how Japan is inviting other nations to participate and share in the results of an R&D project on computer systems of the 1990s.

8. DATA SOURCES AND COMPILATIONS

[1] Buzbee, B.L. and Raveche, H. J.
Report of the Conference on the Forefronts of Large Scale Computational Problems.
This is an informal report on the Conference on the applications of very large scale computing held at the National Bureau of Standards in June, 1984. This report gives a brief summary of the content of papers covering applications in chemistry, medical imaging, materials science, structural analysis, computer generated movies, economics, physics, computational fluid dynamics and also discusses some more general materials.

[2] No Author Given.
Cray Channels.
This magazine focuses on applications of very large scale computing. It is published by Cray Research, Incorporated as a means of illustrating the spectrum of computations which are done on large scale scientific computers.

[3] No Author Given.
Supercomputers.
The Inter. Aerospace Abstracts Data Base 1972 - 1983.
This bibliography contains citations concerning algorithms, architecture and technology of supercomputers. Processing algorithms and architecture, network reliability and inference and problem solving mechanisms are discussed. Applications in laser techniques, nuclear devices, space technology and engineering design and operations are included.

[4] No Author Given.
Supercomputers.
Data Base of the Inter. Infor. Service for the Phy. and Eng. Communities 1975 - 1983.
This bibliography contains citations concerning algorithms, architecture and technology of supercomputers. Fundamental techniques and basic application systems are discussed. Problem solving and inference mechanisms and cooling and packaging systems for supercomputers are presented.

[5] Fuqua, D.
Hearings Before the Committee on Science and Technology.
Record of hearings includes a number of significant position papers and position statements on very high performance computing. Many of the articles here are thoughtful position statements by responsible authorities and in several cases, significant bibliographies are included. This is an important document for those interested in government national policies on large scale computing.

[6] No Author Given.
Proceedings of the IEEE.
Special issue on supercomputers-their impact on science and technology. This special issue of IEEE Transactions includes a number of articles on the numerical applications of very high performance computers. Each of the articles gives an interesting discussion of a particular application area and often includes a very substantial list of technical references.

[7] No Author Given.
Future Systems, State of the Art Report.
Infotech International, Maidenhead, England, 1977.
A collection of papers dealing with computer technology advances, database processors, multimicrosystems, computer market factors, array and associative processor hardware technology and supercomputer architecture. The report includes an annotated bibliography.

[8] No Author Given.
Physics Today.
This is a special issue of the American Institute of Physics publication which gives a number of semi-technical papers on 'doing physics with computers'.

[9] No Author Given.
Physics Today.
This issue contains four articles on application of very high performance computing in the modeling behavior of equilibrium fluids.

[10] No Author Given.
IEEE Spectrum.
Special issue on the next generation. This special issue of IEEE Spectrum contains a list of papers of interest as cultural background for the development of very large scale computing, particularly in the nonnumeric applications.

RESEARCH BRIEFINGS 1984

for the Office of Science and Technology Policy,
the National Science Foundation,
and Selected Federal Departments and Agencies

Committee on Science, Engineering,
and Public Policy

National Academy of Sciences

National Academy of Engineering

Institute of Medicine

NATIONAL ACADEMY PRESS
Washington, D.C. 1984

Report of the Research Briefing Panel on Computer Architecture

Jacob T. Schwartz (*Chairman*), New York University, New York, N.Y.
Richards Adrion, National Science Foundation, Washington, D.C.
Donald Bonstrom, ETA Systems, St. Paul, Minn.
Charles Bostick, National Security Agency, Fort Meade, Md.
John Cavalini, Department of Energy, Washington, D.C.
Robert J. Douglass, Los Alamos National Laboratory, Los Alamos, N. Mex.
W. Daniel Hillis, Thinking Machines Corporation, Waltham, Mass.
Thomas A. Keenan, National Science Foundation, Washington, D.C.
David Kuck, Kuck and Associates, Champaign, Ill.
H. T. Kung, Carnegie-Mellon University, Pittsburgh, Pa.
Thomas Nash, Fermi National Accelerator Laboratory, Batavia, Ill.
James Pomerene, IBM Corporation, Yorktown Heights, N.Y.
John Riganati, National Bureau of Standards, Gaithersburg, Md.
Arnold Reisman, Microelectronics Center of North Carolina, Research Triangle Park, N.C.
Paul Schneck, Office of Naval Research, Arlington, Va.
Stephen Squires, DARPA/IPTO, Arlington, Va.
Kenneth Wallgren, NASA, Washington, D.C.

Staff

Mark Weiser, *Rapporteur*, University of Maryland, College Park, Md.
Gary E. Clark, *Staff Director*, Computer Science and Technology Board
Rosena A. Ricks, *Administrative Secretary*, Computer Science and Technology Board
Allan R. Hoffman, *Executive Director*, Committee on Science, Engineering, and Public Policy

Report of the Research Briefing Panel on Computer Architecture

I. INTRODUCTION

The market for supercomputers is only a small part of the total computer market. However, it represents a critical cutting edge for the advance of computer technology and provides machines that are critical to technological leadership in a whole range of areas varying from aircraft design to basic chemistry. For this reason it is vital that the programs that achieved U.S. supercomputer leadership be continued and that new programs to maintain that leadership be initiated.

Rapidly developing VLSI technology has created the basis for major advances in superspeed computer performance. Maximum attainable computation rates, which are approaching 1,000 MIPS (millions of instructions per second) for the largest vector supercomputers, will rise over the next decade through 20,000 MIPS, and possibly to levels of 100,000 MIPS and more. Attaining such extreme speeds will depend on the effective use of parallelism, i.e., on the development of computers that can execute many hundreds, thousands, or tens of thousands of instructions simultaneously and of software that can effectively orchestrate many simultaneous streams of computation. The ability to explore this essential new opportunity has been opened up by the amazing reductions in circuit costs attained during the past decade.

This document will review the current supercomputer situation and will predict developments over the next 5 to 10 years. The following major areas will be considered:

1. the *chip and mass storage technologies* on which U.S. primacy in the computer area, and its ability to advance, ultimately rest;
2. the *computer architectures*, both novel and relatively conventional, likely to be important for large-scale computing over the next decade;
3. the potentially significant role of *special-purpose computers*, which may be of particular importance for certain types of research (e.g., computer vision, speech and signal processing, scientific data reconstruction and reduction), where dedicated equipment and very high performance, but relatively inexpensive equipment, are all essential;
4. the extent to which *software technology* can and must support the move to large-scale parallel computing;
5. the developments in *communication tech-*

nology that will define the ways in which supercomputers can be made available to the U.S. research community;

6. the *efforts abroad* that might undercut the dominant U.S. role in high-performance computing;
7. the relationship of supercomputing to the broader field of *artificial intelligence*; and
8. the bottom-line technical policy question of how to get the U.S. research and industrial communities to *move forward in rapid and effective cooperation*.

II. CHIP TECHNOLOGY

The astonishing exponential growth of microchip performance is both the main driving force for the computer industry and the technical basis for expected parallel supercomputer developments. Though many circuit families of differing speeds and cooling requirements play a role in this growth, the following remarks will concentrate on just one family, the so-called field-effect transistors (FETs), which generate 40 times less heat than the faster bipolar chips, a very significant advantage as the number of circuits per chip increases steadily. State-of-the-art FET chips as of 1984 are represented by the 256K memory chip (512,000 devices), just now coming into mass production, and the roughly 100,000-device Motorola 68000 microprocessor. The 5 to 1 circuit density advantage of memory chips results from their very regular internal geometric structure and highly optimized design. Since 1960 these circuit densities have doubled roughly every 18 months, and this growth is expected to continue, with memory and microprocessor densities remaining in a similar proportion. By 1988 this should make 2-million-bit/chip memories available and could make million-transistor single-chip processors possible. As circuit sizes shrink, speeds will rise, so by 1988 the cycle-time of single-chip FET processors should have fallen from the 80 nanosec now typical to 30 nanosec for chips operated at room temperature and 15 nanosec for chips operated cryogenically. A typical application of this technology might be a single-chip, 64-bit scientific processor with full floating-point capability, on-chip memory management, internal overlap of instruction execution, 32 Kbytes of on-chip cache memory, and the capability of performing 20-40 MIPS. Processors tailored for somewhat faster execution of higher-level languages such as LISP and special-purpose computing devices like those described in Section V below (e.g., specialized vision-oriented or switching chips) all can be supported by this same technology.

The use of new materials such as gallium arsenide and other techniques for attaining higher electron mobility should allow perhaps 10 times higher performance, but they require significantly more difficult process and materials technology. The Japanese HEMT and the U.S. VHSIC are examples of such projects. Josephson circuits, which operate at cryogenic temperatures, and 3-D chips, which try to build up circuits in the third dimension rather than using the essentially planar structures now common, are not likely to be useful technologies in the near future.

Wafer-scale integration, including full and partial wafer techniques, have been considered for reducing the costs associated with the currently used wafer-dicing and reassembly technology. This approach faces severe chip yield problems, and one must reckon with the unfavorable signal-propagation characteristics of long on-chip conductors. For this reason, a packaging technique like that used in the IBM Thermal Conduction Modules, i.e., bonding individual chips to a multilayer ceramic substrate, may dominate for high-performance devices. Assuming that 200-400 chips are bonded to a single 6″ × 6″ ceramic board of this type, two such boards might by 1990 produce a 64-processor, 1,200 MIPS peak-rate parallel computer at a manufacturing cost between $50 thousand and $100 thousand.

Attaining the chip performances described in the preceding paragraphs will require attention to a challenging mass of supporting technologies, including special materials development, x-ray, E-beam, and ion lithography, as well as radiation damage removal

methods. This work is moving rapidly in the United States and at least as rapidly in Japan, where experimental and fabrication facilities sometimes bettering the best U.S. facilities have been put in place. Japanese work in this critical area is spurred by a particularly clear national determination to succeed, resulting in a willingness to try approaches that U.S. workers tend to avoid because of initially unclear market perspectives. In this determination, combined with the dedicated large teams and engineering excellence brought to bear, lies the technological roots of the Japanese computer industry's rapid advance.

III. DISK TECHNOLOGIES

For supercomputers to be used effectively, they must be supplied as balanced systems, with input-output (I/O) bandwidth, external storage, displays, and other capabilities all appropriately matched to the expected two- or three-orders of magnitude increase in computation rates. Concern has been expressed about the ability of storage disks, a key element of this "peripheral" technology, to keep up with such rapid speed increases. The workhorse disk for current supercomputer systems is still the CDC Model 819 disk, a rotating storage device dating from 1970, with a capacity of 600 million bytes and a 5-Mbyte-per-second transfer rate. Two newer disks just now becoming available provide roughly double this capacity and storage rate, an exceedingly modest increase compared to far more rapid computation-rate speedups. Other more ambitious magnetic disk alternatives include the possible use of synchronized spindles or multiple heads in parallel. Although no such higher performance products have been announced (probably because of less-than-satisfactory past market experience with very-high-performance disks), products of this type may appear as derivatives of disk and head technology developed for larger, more lucrative markets.

Against this background, the development of a new optical disk storing 10^{11} bytes, now in progress under NASA and USAF (RADC) sponsorship, appears extremely significant. This disk, a first version of which should be available in 1987, promises to support a 10^8 byte/sec transfer, much better matched to anticipated supercomputer requirements than anything currently available. Access time, i.e., the time needed to start the inflow of a randomly selected data record, is anticipated to be 50-100 milliseconds. The development of this potentially important device ought to be pushed rapidly to prevent an initially limited market from delaying its appearance.

IV. CURRENT AND NEW COMPUTER ARCHITECTURES

During the next 5 years, new types of machines based on large-scale parallelism, having the ability to perform thousands or even tens of thousands of arithmetic operations simultaneously, will begin to appear alongside high-speed sequential and vector computers that dominate today's high-end computer market. At a rate determined by their performance, their cost, and their suitability for various applications, these new machines may progressively supplant present designs in some market sectors. Though our panel did not make any precise predictions concerning the rate at which this development will unfold over the next decade, we can describe the main architectural considerations relevant to this process.

High-speed single-thread machines keep programming relatively straightforward by behaving externally as if they execute instructions serially, even though internally a small number of operations (5-20) are actually being handled simultaneously. Machines of this design are the staple of present medium-high performance (4-20 MIPS) computing. It is technically possible to raise the performance of these machines to the 80-150 MIPS range with current technologies and to 250 MIPS or somewhat more using more advanced technology such as that represented by GaAs. Performance improvements beyond this will

depend on the parallel use of multiple high-speed serial processors. Very-high-speed serial machines and multiprocessor configurations involving small numbers of such computers are bound to retain major, perhaps dominant, economic importance for the next 5-10 years, especially for commercial applications. In the supercomputer environment, fast (4-10 MIPS) serial computers will be the single-user "workstations" that supply researchers with their software development tools and much of their graphic and data display environments.

Vector supercomputers, like the present U.S. Cyber 205, CRAY XMP, and the Japanese Hitachi S-810/20 and Fujitsu VP-200, can be regarded as simple parallel machines specialized for the handling of computations characterized by particularly regular patterns of data motion. Such regularities typify certain important numerical scientific codes, and hence fast vector machines are generally designed as "scientific number crunchers" equipped with the highest-speed floating-point operation units. The current peak performance of these machines is in the hundreds of MIPS. Single-processor machines of this type could achieve operation rates of 250-500 MIPS with the best present technology and perhaps twice this using GaAs technology. Multiprocessor configurations of vector processors, a type of design under active consideration by several supercomputer manufacturers, will probably attain peak performances of roughly 10,000 MIPS* using 16-32 processors within 5 years, rising to 25,000 MIPS within the coming decade. At such performance levels, they can be expected to retain a major role, and perhaps a dominant role, in bread-and-butter scientific number crunching through much of the decade. Note, however, that computers of this design can exploit parallelism only in innermost program loops; and for programs that branch or address memory in irregular patterns, the actual efficiency typically decreases to 15-20 percent of rated peak performance, e.g., for such problems a 4,000 MIPS actual computation rate is all that can be expected from a machine nominally rated at 25,000 MIPS.

New parallel supercomputer designs, many of which are still highly experimental, are at once the focus of much current research and the one area (of design, as distinct from technology) in which sudden surprising advances might occur. Many of these designs promise large increases in computational power without significant redesign of their components, i.e. they are "scalable." With such architectures it may be possible to increase machine speeds by factors of 10 or even 100 by increasing the amount of parallelism used, without major redesign but at a proportional increase in cost. Once suitable software has been developed for machines of this type, a major research undertaking, the homogeneity of their structure should make it possible to move applications to much larger-scale parallel machines with minimal software change. Moreover, scalable parallel machines have major potential advantages in a VLSI environment because they are built from large numbers of identical parts that can be mass produced efficiently and because the design cost can be amortized over multiple configurations differing only in size. For all these reasons, we review this class of computers at length and attempt to clarify the technical options available for parallel supercomputer development by listing various principal design alternatives characterizing the design of such machines.

Most of the more than 70 parallel machines that have been proposed by university and

*The performance of high-speed scientific computers is often rated not in terms of the total rate at which instructions of all types are performed but only counting that part of the total activity devoted to the execution of "floating-point" arithmetic operations: MFLOPS (millions of floating-point instructions per second) rather than MIPS (millions of instructions per second). The ratio between these two measures is typically taken to be 1/3, but since this ratio is highly application dependent, we have chosen to sidestep this distinction and to state all crude performance figures in MIPS.

industrial groups here and abroad occupy easily characterized positions in the "design space" defined by the following alternatives:

1. *SIMD versus MIMD.* An SIMD (single-instruction multiple-data) machine consists of multiple processors (or of a few very fast pipelined processing elements), all of which are fed synchronously by a single instruction stream but which operate on independent streams of data. Vector machines like the CRAY I or Hitachi 810, and synchronous array machines like the ILLIAC IV and NASA/Goodyear MPP, typify this class. An MIMD (multiple-instruction multiple-data) computer consists of multiple processors (or fast time-sliced processing elements) driven by independent instruction streams and capable of branching independently but also able to pass data and synchronization signals between processors, perhaps via a shared memory. The Denelcor HEP, Cal Tech Homogeneous Hypercube, and numerous other machines belong to this class. SIMD machines are generally more efficient for applications characterized by very regular patterns of processing, in part because their "lockstep" mode of operation shares a single unit of control hardware among many data-processing elements and eliminates much of the message-passing overhead needed when all the processors of an MIMD machine must proceed through a computation in close synchrony. On the other hand, it is hard to make SIMD computers deal effectively with highly "branched" code, since the process of branching in such computers is normally imitated by temporarily disabling some of the processors in the parallel array, and thus several successive branches tend to reduce effective computational power drastically. MIMD processing arrays deal more robustly with complex branched code, since they are capable of executing not only innermost loops, but even complex branch-filled loops, in parallel.
2. *Visible versus "under the covers" physical communication pattern.* In all parallel machines some portion of the computation involves communication among the individual processing elements. From the physical point of view such communication is allowed in only a few specific patterns defined by the hardware. For example, the processors may be arranged in an obvious two-dimensional grid with each processor connected to its north, south, east, and west neighbor. A single operation on a machine wired in this pattern can only send a number from each processor to one of these neighbors. Proposed connection patterns for parallel machines include rings, n-dimensional cubes, binary trees, and various less obvious but much more powerful communication networks developed by communications researchers.

   Once having chosen a communication net appropriate for the computational structure being developed, a parallel machine designer can choose either to make the structure of this net visible to the machine's user or to hide these details behind a set of programmed conventions that makes it appear that any processor can communicate directly with any other.

   Both alternatives have arguable advantages. On the one hand, if a machine's actual communication pattern is well matched to an application and is made available to the programmer, it can be exploited to gain speed. Examples are the use of two-dimensional grid patterns for image processing and the use of various powerful communication patterns for optimal implementation of the Fast Fourier Transform. On the other hand, machines that relieve their programmer from direct concern for the pattern in which messages must flow

between processors are easier to program for a wide range of problems, particularly problems characterized by highly variable patterns of communication. Another potential advantage of machines in this latter class is that the communication paths that they use, being hidden, can be varied dynamically, e.g., to bypass faulty components.

3. *Coarse-grained versus fine-grained designs.* In any parallel computer with multiple processing elements there is a cost tradeoff between the number and the size of the processors. The approach that stays closest to established single-stream processor technology is to use as few as possible of the largest available processors. The opposite approach is to achieve as much parallelism as possible by using a large number of small elementary processors, which can be single-chip microprocessors chosen to exploit the cost-effectiveness of a mass-produced item or, in the limit, can be a single-bit processor used in very large numbers. The decision to use a relatively small number of very-high-speed processors is justifiable by the argument that even largely parallelizable code will have significant serial sections, so high performance in serial code will be important for sustained overall high performance. Its polar opposite, the use of many single-bit processors, can be justified by a desire to push parallel execution to its limit, and by the surprising level of arithmetic power per square millimeter of silicon that one-bit processors attain. Perhaps the most important issue here is one of programming style. Since ordinary serial machines are coarse-grained, the technology for programming coarse-grained machines is understood best. Thus, it is plausible to expect a FORTRAN compiler to optimize code that keeps, say, 16 processing units busy, but not 16,000. On the other hand, if an algorithm is written with parallel processing in mind from the start, it may divide naturally into numerous processes, e.g., the processing of a 1,000 by 1,000 image in which relatively simple, low-precision algorithms must be applied to individual pixels or small pixel groups may fit most naturally on a million processor machine.

4. *Specialization of architecture to an anticipated application or programming style.* Scientific number crunchers are optimized for arithmetic operations on floating-point numbers, while processors intended for more varied symbolic applications must be optimized to handle irregular memory references and branching patterns, fast subroutine transitions, and fast context switching. They must provide run-time type checking and must support efficient execution monitoring during program debugging. Thus in designing a "scientific" processor one will emphasize certain types of hardware that are of little use in a "symbolic" machine, and vice versa. Various designs for large-scale "dataflow" computers, which are specialized in a different direction, illustrate this same general point. Dataflow programmming is organized around a set of fully parallel concepts. A program is regarded as an assemblage of "operators," all poised to fire as soon as they receive all their logical inputs. As soon as it fires, each such operator transmits its output to further operators, for which it becomes one of several possible inputs.

   The innermost interpretation cycle for such a language consists of various operand-routing, instruction-dispatching, and instruction-execution activities, many of which can be overlapped. As compared to a general-purpose parallel assemblage, a cleverly designed layout of dataflow operation nodes in

memory and effective instruction dispatch hardware can probably double or triple the efficiency with which a dataflow language executes, while adding only minimally to the total hardware cost of a parallel system. These remarks typify the speed gains achievable by the architectural specialization of parallel machines.

V. SPECIAL-PURPOSE COMPUTING DEVICES

Extremely high computation rates can be attained efficiently by tailoring electronic hardware to algorithms of special importance (e.g., the Fast Fourier Transform) or to the requirements of a particular computer-intensive problem. Devices of this sort are of particular importance to diverse cutting-edge research efforts and to applications where dedicated equipment is essential but dedicated use of multimillion dollar general-purpose computers such as a large vector supercomputer would be prohibitively expensive.

These special-purpose computing devices point to a class of computing devices highly optimized to particular applications, which might be made possible by hypothetical future advances in compiling technology that allow automatic generation of whatever computing device is best suited for any specified application. Though this is far beyond present capabilities, manually designed special-purpose computing systems are of growing significance to defense uses, such as signal processing and image analysis, to high-energy-physics data acquisition and reduction, and to other areas.

It is steadily becoming easier to realize such systems, either by combining relatively standardized high-performance modules into specially wired configurations or by custom gate-array or VLSI designs. Special systems produced in this way tend to reflect end-user performance requirements closely, and hence they uncover important subproblems of the general supercomputer design problem quickly. Moreover, because of the emphasis on cost-effectiveness typically characterizing their design, they encourage industry to improve the effectiveness of more general computing engines supplied commercially. Also, though machines of this sort sometimes involve approaches that address only one specific problem, others affect larger classes of problems in an extremely cost-effective way, making them desirable attachments for general-purpose supercomputer systems. Encouraging a steady flow of such designs, which sometimes advance the effective frontier of computing in the application areas they address, can strengthen the U.S. competitive position significantly.

The sum of these special design efforts has become an important component of computer research that needs to be explicitly recognized and systematically cultivated. For this, an improved infrastructure for rapid and effective development of special-purpose systems is a prime necessity. In particular, projects motivated by special problems need a basic kit of standard modules out of which optimized architectures can be built. This should include low-cost single-board computers, high-density multiport memories, and high-speed switching modules that allow large communications networks tailored to specific applications to be fabricated rapidly. Some of these items are beginning to appear commercially, but further encouragement to industry and systematic attention to standardization issues are appropriate.

To increase the speed with which prototype special-purpose computational designs can be developed for significant experiments, system-level tools of the same quality as the best current VLSI design tools need to be developed. A focused program to develop such tools and make them available to universities, laboratories, and private-sector organizations appears to be an effective way to strengthen the work of special-system designers. Such tools would also facilitate the transfer of successful prototypes into indus-

trially supported products, which is as important for special high-performance computational designs as for general-purpose supercomputer development.

VI. THE ROLE OF COMMUNICATION TECHNOLOGY

Supercomputers communicate three ways: internally, locally (in the same building), and outside (to machines elsewhere in the world). Existing technology for internal connections uses highly parallel buses and is limited primarily by wire length. Existing nearby machine telecommunications technology runs at a maximum of about 200 Mbit/sec, which is faster than most individual computers can accept information. Current outside connections are limited by available nationwide channels to 3 Mbit/sec, and 50 Kbit/sec is more typical.

In the next 5-10 years it is unlikely that the basic internal communication technology will change much. However, it is possible that within 20 years optical connections or wafer-scale integrated circuit technology will have significantly changed the way computers communicate internally.

Local interconnections will be affected by optical technology much sooner. It is possible that within the next 5-10 years commercial optical local networks running at a gigabit/sec rate will become available. This rate will not be accessible to any individual machine but will permit many supercomputers to share a relatively inexpensive network and exchange data at their internal bus rates.

NBS, DARPA, and cooperating international standards organizations have made large-scale networks possible by specifying protocols and standards that promote interoperability of heterogeneous equipment from different manufacturers. However, nationwide network speeds are currently inadequate for the amount of data consumed by supercomputers. Over the next 5 years the NSF ScienceNet project should be in place with high-speed satellite connections linking the major research universities. This will be a significant help to providing access to supercomputers.

Finally, we note that communication security on these nets may become an important issue, and this area deserves careful and early attention.

VII. ALGORITHMIC AND SOFTWARE ISSUES

Effective parallel processing will require a fundamental understanding of computational models, algorithms, programming languages, optimization techniques, processor-memory-I/O relationships, and interfaces to existing computing systems and end users. Though research on parallel algorithms is expanding, the lack of large-scale parallel machines for experimentation orients current work toward the study of individual algorithms rather than on practical applications, systems questions, or large software structures. Moreover, systematic work on large algorithm libraries, such as would quickly develop around major parallel computing installations, is hardly possible in the present atmosphere of largely theoretical studies. Also, although operating systems for parallel machines raise many subtle problems, it is difficult to address these except by working with real parallel systems.

Parallel programming is much more complicated than sequential programming because it deals with simultaneously occurring events involving potentially complex (time-dependent) interactions, so that elusive and nonrepeatable bugs can occur. For this reason, the usability of large parallel machines will be bound up with the development of programming languages that facilitate the expression of parallelism and the development of sophisticated optimizing compilers that improve the execution efficiency of such languages and find implicit parallelism where it is not explicitly expressed.

Hence, although initial experience (e.g., work on the HEP at the Los Alamos Labora-

tory) indicates that modest extensions to conventional languages will allow important applications to run with high efficiency on certain types of parallel computers, many new programming language concepts need to be investigated as ways for easing the further exploitation of parallelism. These include languages with vector and set operations (e.g., APL), dataflow languages, applicative languages, languages such as parallel LISP and PROLOG that emphasize recursive branching and backtracking processes, and languages in which communicating sequential processes play a central role. It is important both to encourage theoretical study and development of all these classes of languages and also to ensure that they become available for extensive experimental use as soon as substantial parallel computers suitable for them are constructed.

The present state-of-the-art software to support high-speed parallel computation is also represented by vectorizing compilers and more general parallelizing compilers capable of finding certain opportunities for parallelism automatically by analyzing ostensibly serial code. Both Cray Research and Fujitsu Ltd. have produced such compilers for their vector supercomputers, as have CDC, and U.S. research projects at the University of Illinois and Rice University.

While very useful, these systems are not yet able to automatically translate every standard sequential FORTRAN program into a good parallel program. Indeed, this will remain beyond the capabilities of automatic systems in the foreseeable future because the best parallelization usually requires substantial comprehension of the underlying problem domain. However, these systems do provide invaluable aids to the programmer because they permit him to write parallel programs cleanly in a high-level language without cluttering the language with special-purpose constructs.

Present automatic code vectorization techniques need to be evolved to handle programming for parallel multiprocessor supercomputers and to apply to whatever new parallel languages appear most important. Many of today's vectorization techniques are directly applicable; others will need to be adapted substantially or new methods invented. Program development environments tailored to the special requirements of new parallel systems and languages also need to be developed, especially to aid in the debugging of parallel programs. Graphics systems making it easy to deal with the complex outputs that parallel supercomputers can generate will also be needed.

Complex software developments cannot proceed far in a purely "paper mode." For them to flourish, real and relatively large-scale experiments with functioning hardware are essential. Thus, to accelerate the very challenging software research outlined in the preceding paragraphs, it is essential to deploy or provide remote access to experimental parallel machines as soon as they become available. Unconfining access is essential to the major systems software research and development efforts that will be required to ensure effective use of these new machines.

VIII. EFFORTS ABROAD

U.S. supercomputer developments must be appreciated in the context of efforts abroad, primarily those of the powerful and fast-moving Japanese computer industry, and also relative to an accelerating European effort. Our panel's conclusions concerning the Japanese program are as follows. The largest Japanese vector supercomputers (Hitachi S-810/820 and Fujitsu VP-200) have attained parity with, but not surpassed, corresponding U.S. machines such as the CDC Cyber 205 and the CRAY XMP. Japan's stated national commitment to advance in this area, the great engineering strengths of Japanese industry, and the leading position that Japan has won in certain lines of high-speed semiconductor component production will make for a very closely run race in this area. Nonetheless, the United States can at least expect to maintain

parity in vector supercomputer production, and moreover in various rapidly expanding U.S. supercomputer use-oriented educational and research programs, e.g., the newly announced NSF Supercomputer Access Program, will help preserve U.S. leadership in vector supercomputer applications development.

Competition in regard to highly parallel supercomputers of newer experimental designs favors the United States even more strongly. The United States can expect to retain world leadership in this area, provided that agency support focuses on the most robust new designs and that proper bridges are built between university-centered innovation and the organizational and manufacturing capabilities of the U.S. computer industry. Without such links, it is feared that competition in this area will find various relatively small U.S. firms, including some recent start-ups, forced to compete with the extremely impressive technological capabilities of such major Japanese firms as Fujitsu, Hitachi, and NEC. It is noted in this connection that some Japanese firms have evinced interest in exchanging technology and development support for the results of U.S. conceptual research.

Software construction is also an area in which Japanese strengths, through not yet widely recognized in the United States, are already impressive, as shown, for example, by the first-rate vectorizing compilers available with the new Hitachi and Fujitsu vector supercomputers. Though the technology involved derives almost wholly from U.S. work, a recent study of Japanese software engineering practice, done by the JTECH Panel on Computer Science under the sponsorship of the U.S. Department of Commerce, concludes that in the production of standard software systems Japan may already be ahead of the United States and may be getting further ahead. For example, the average software production per employee in the 2,000-employee Toshiba software factory (which produces complex distributed systems process control software for power plants, steel mills, flight guidance, and so forth) is reported at 2,000 lines of high-level code per month, as compared to a typical U.S. figure of 300 lines per month. Error rates of 0.3 errors per 1,000 lines are achieved, as compared to U.S. rates 10 times as high, which may provide Japanese vendors the competitive advantage of marketing software with a 10-year warranty. Much higher rates of software module re-use than characterize U.S. practice are attained. The cited JTECH Panel report ascribes these significant successes to the systematic application of disciplined coding techniques and to source code control methods, quite familiar in the United States but much less systematically applied, and also to the sense of responsibility for a product that pervades Japanese society and to the factory discipline that Japanese management is able to apply to software production. This Japanese prowess in an area far more central to the computer industry than supercomputer production can potentially erode U.S. systems sales significantly and calls for further prompt analysis and appropriate response.

In spite of the intense international attention and publicity that has been focused on Japan's "Fifth Generation" computer effort, it is not clear to our panel that the ambitious goals set for it are matched by any entirely persuasive technical idea of how they can be reached. Nevertheless, it must be understood that even if not fully successful the kind of ambitious long-range research represented by the Fifth Generation project will generate significant new engineering, architectural, and software strengths for Japan. It must also be understood that Japanese universities and firms have mounted many parallel-computer explorations smaller and less publicized than the Fifth Generation effort and that cumulatively these will position Japan to move forward rapidly with new parallel machine designs. It is quite possible that these new Japanese designs will emerge immediately after the current wave of U.S. effort has

determined what lines of development are most promising.

We note also that it is clearly in the U.S. interest to build up a much larger body of computer specialists familiar with Japanese work and able to read Japanese technical literature. As long as tens of thousands of Japanese researchers follow the U.S. literature diligently while only a handful of U.S. computer scientists can deal with Japanese, the main flow of technical information will clearly run from the United States to Japan.

The European "Esprit" program does not appear to confront the United States with competition as severe as that seen in Japan. A review of the long list of projects to be supported under this program suggests that it will serve to strengthen European work in a bróad range of technological areas, but concentrated work apt to result in supercomputer developments that will create major competition for the United States does not appear likely. Nevertheless, this European work needs to be followed closely.

The British response to the Fifth Generation challenge is the "Alvey" program, which is similar in its goals to the U.S. MCC research consortium now established in Austin, Tex. The four technologies making up Alvey are VLSI, software engineering, AI, and man-machine interfaces. Extensive university-industry cooperation is planned, and U.S. companies are welcome to participate as long as the research is done in Britain. Although the Alvey focus on supercomputer issues is not yet clear, it is important that this potentially significant national effort be followed closely.

IX. THE ARTIFICIAL INTELLIGENCE ISSUE

Construction of computerized artificial intelligences can justly be regarded as one of the most central long-term goals of computer science (and even of science generally), and it deserves tenaciously sustained support. The supercomputer developments now in progress will undoubtedly advance artificial intelligence research by putting new technical means at its disposal. An initially modest but growing degree of specialization in computer architectures can be expected to develop in response to the requirements of this field. Though the algorithmic structure of artificial intelligence computation is much less mature and stable than that of other areas of computation (e.g., numerical scientific computation or computer algebra), it is clear that work in artificial intelligence can be expected to emphasize random patterns of memory reference, various exact and approximate matching processes, computations with short quantities rather than multidigit floating-point quantities, and languages such as LISP whose execution can be accelerated (perhaps tripled) by suitable hardware. Moreover, the first stages of data input for subfields of artificial intelligence such as image and speech processing will require various special-purpose computational devices of the kind discussed in Section V, while exploration of various programming styles currently of interest to workers in artificial intelligence, e.g., fast interpretation of large systems of production rules, high-speed formula unification in logic, or neutral-net simulation, may also justify specialized hardware. The scientific computation and artificial intelligence communities favor distinct programming languages, emphasize different aspects of system performance, and prefer different operating systems, making it difficult for them to be served effectively by a single operational environment. Nevertheless, it remains difficult to derive any specific supercomputer designs from the requirements of artificial intelligence research, and in this sense the division of supercomputers into two distinct genuses of "scientific machines" and "artificial intelligence machines" is artificial. Perhaps it is best to say that artificial intelligence, as a young field in which approaches are still evolving, requires computers that can deal

with rapidly shifting patterns of computation and communication, with great flexibility and at very high speed.

X. HOW TO MOVE FORWARD

Several major federal initiatives aimed at strengthening U.S. supercomputer-related capabilities have been launched and, if pursued along the strongest technical lines and with due regard for the importance of early and effective industrial involvement, can maintain U.S. computer preeminence. Principal among these are the DARPA Strategic Computing Program, the expanded supercomputer research programs of the Department of Energy, and NSF's Supercomputer Access Program. The steady support that NSF has provided for many, indeed most, of the university groups whose work has defined the principal supercomputer architectural alternatives now available has been of fundamental importance and is planned to continue and even to increase modestly.

Since DARPA's architecture program comprises the largest single federal project addressed to supercomputing, its activities will be broadest and most crucial. Work in many areas, including underlying technologies, improved design automation tools, fabrication means for both VLSI and broad-level special systems construction, conceptual exploration of new special general-purpose parallel computer systems, various small-scale prototyping efforts, and a few critical large-scale prototyping efforts, will all be sponsored.

Our panel is supportive of DARPA's intent to organize the selection of the culminating major efforts of this multiyear program via an orderly series of phases, beginning with conceptual studies aimed at the discovery of significant new computer architectures, progressing to the exploration of those architectures by simulation and experimental software development, then to small-scale hardware prototyping where justified, and finally, in the most promising cases, to the construction of large high-performance prototypes. We note that for these large culminating prototypes, research success, and even more the ability to move successful research results into broad U.S. use before foreign competition follows suit, will clearly be dependent on serious industrial involvement. The proper use of academic resources is in research and small-prototype development; the realization of large systems requires the skills of proven engineering organizations with established track records. For these reasons, we recommend that both DARPA and DOE encourage or require substantial industrial involvement (by either established or start-up companies) in the development of each experimental computer system chosen for large-scale prototyping. We understand that it is DARPA's intent to encourage this type of industrial involvement and we support this decision. Ideally, this should involve a substantial commitment of capital and crucial technical skills by the company (or companies) cooperating in such a development, to ensure that careful and independent private-sector technical judgment concerning project soundness is brought to bear in each such case. Moreover, to strengthen university-industry links and overcome what might otherwise be substantial familiarization delays in moving promising designs into large-scale prototyping, it is recommended that both DARPA and DOE encourage industrial involvement even in the initial smaller-scale stages of parallel computer prototyping. Early industrial involvement in special-purpose computer developments is also desirable.

Though considerably smaller than DARPA's effort, the DOE supercomputer program has maintained a significant focus in areas complementary to, and synergistic with, the DARPA effort. Historically, DOE has been the leading agency for the applications of large scientific computing and has maintained

particularly close relationships with the supercomputing community. Since optimization of computer architectures for floating-point computation and for the regular patterns of data manipulation characterizing many numerical applications may demand significantly different designs than are optimal for nonnumerical applications, it is important to maintain a strong program of architectural research and development with a clear scientific-computation focus. DOE's strong tradition and high level of expertise in scientific supercomputing make it appropriate for DOE to play a significant funding role for research on advanced computer architectures aimed at high-performance numerical computing, and acceleration of its program in this area is desirable. DOE is also in the process of significantly expanding its supercomputer access program to include scientists in High Energy and Nuclear Physics, Basic Energy Sciences, and Biological and Environmental Research. This expansion will contribute substantially to solving the national supercomputer access problem. In addition, DOE has acquired several commercial parallel computers and is making them available to many researchers. These DOE programs have played and will continue to play a very important role in maintaining overall supercomputer application capabilities.

As currently defined, NSF efforts will concentrate on making supercomputer capabilities broadly available to the U.S. scientific community, including but not restricted to computer scientists. Though initially this program will concentrate on providing access to existing vector supercomputers, our panel sees the NSF Access Program as having broad continuing implications and recommends that NSF plan to make all major experimental parallel computers nationally accessible as soon as feasible. Efforts of this sort, which serve to build supercomputer-use capabilities, strengthen the U.S. position by exploiting one of our most characteristic advantages, the great breadth and diversity of the U.S. computational community. However, the resources of this community will be heavily taxed by the deep algorithmic and software problems associated with the new classes of machines that will be developed over the next decade. Training of an expanded population of workers in this field must therefore be seen as an essential component of a balanced national effort.

The NSF Supercomputer Access Program is well calculated to serve the requirements of the large U.S. scientific community for whom computers are a tool for other research goals. However, in view of the great importance that NSF basic computer science support has had in the development of new architectural alternatives and the related computer science until now and in view of the excellent technical judgment with regard to hardware and software that NSF can bring to bear, the NSF programs should also be structured to support major software and applications development for the machines to which access is provided. Moreover, to avoid dissipating the attention of research groups among overly many new machines, all the agencies sponsoring supercomputer research and development should plan to support groups wishing to develop algorithms and software for experimental hardware systems produced commercially or by other universities. Such participating research groups should be guaranteed adequately "privileged" remote access, e.g., the right to take full control of a system for scheduled blocks of time to allow exploration of operating system changes in the light of new research ideas. In some cases, it will also be appropriate to disseminate smaller-scale versions of major experimental systems to several locations.

The need for quantitative measurements of progress and quantitative characterization of the advantages of alternative architectural approaches underlies all the areas of research we have discussed. It is therefore appropri-

ate to encourage joint industry-university-government development of tools and techniques to measure progress in software, architecture, and algorithms for supercomputing systems.

XI. CONCLUSION

The U.S. national interest in the rapidly moving area of computer architecture should be as follows:

1. To ensure that all designs likely to succeed in any major way are explored actively. We must guard against technological surprise.
2. To get successful, or potentially successful, designs into the hands of large user and software development groups quickly, so as to accelerate the development of the necessary software. We must maintain system and software leadership.
3. To accelerate the transition of successful designs into commercial production, preferably by involving strong industrial groups in their development from the start.

It is quite important for U.S. activities aimed at maintaining awareness of Japanese technical developments through suitable translations be accelerated, possibly through programs undertaken by NSF or DARPA.

Finally, since the U.S. faces rapidly mounting competition in this area from both Japan and Europe, it is important that vigorously conceived programs, including those now planned by the major agencies involved, move forward without delay.

REPORT OF THE FEDERAL COORDINATING COUNCIL ON SCIENCE, ENGINEERING, AND TECHNOLOGY PANEL ON ADVANCED COMPUTER RESEARCH IN THE FEDERAL GOVERNMENT

JUNE 1985

REPORT OF THE FEDERAL COORDINATING COUNCIL ON SCIENCE, ENGINEERING, AND TECHNOLOGY PANEL ON ADVANCED COMPUTER RESEARCH IN THE FEDERAL GOVERNMENT

JUNE 1985

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

PREFACE

In the spring of 1983, the Office of Science and Technology Policy (OSTP) formed three interagency panels under the Federal Coordinating Council on Science, Engineering, and Technology (FCCSET) to examine critical emerging issues in the computer field. Two of the FCCSET panels were chaired by the Department of Energy (DOE); these panels focused on the issues of procurement and access to high-speed numerical machines known as supercomputers. Reports covering these two areas were prepared for OSTP by the end of 1983 and distributed widely throughout the government. The two panels were later merged into one to consider the follow-on issues in supercomputer procurement and access. The third FCCSET panel is chaired by the Department of Defense (DOD) and its charter is to stimulate the exchange of information within the government on high-performance symbolic computing and artificial intelligence (AI).

During the course of its initial deliberations, it became clear that interest in information exchange extended beyond symbolic processing and AI. Questions arose concerning the architectural and computational differences between symbolic and numerical computing and the requirements of very high-performance machines that might be able to handle both types of computations effectively. As the individual agencies presented their current and planned research programs, the relationship between these two areas of research became a topic of discussion in its own right. Consequently, it was decided to broaden the charter of the panel beyond symbolic processing and artificial intelligence and to include federal research efforts in very high-performance scientific and numerical computing and Advanced Computer Research in general.

This document constitutes the first report of the FCCSET Computer Research Coordination Panel. It provides an overview of current and planned federally sponsored research activities in Advanced Computer Research and Very High-Performance Computing, in particular. It summarizes existing activities funded by individual agencies and by more than one agency, identifies the FY 1983 to FY 1985 funding, and presents a set of findings and recommendations for further consideration.

Within the Government, the primary supporters of basic computer science research are the National Science Foundation (NSF), Department of Energy (DOE), and the Defense Advanced Research Projects Agency (DARPA). The National Aeronautics and Space Administration (NASA), and the military services, while conducting limited basic research programs, complement these efforts with mission-specific investigations. In addition, the Department of Commerce, through the National Bureau of Standards (DOC/NBS), is concentrating on standards, metrics, and benchmarks, while the National Security Agency (NSA) and the Central Intelligence Agency (CIA) are studying the intelligence aspects of the technology.

In the course of its initial discussions, the panel focused on the Federal role in advancing the state-of-the-art in computer science research. However, the sheer size and diversity of the field made categorization, much less coordination, a challenge. It also became apparent that coordination was most appropriate and desirable in the subset of research which the panel identified as Very High-Performance Computing; each of the participating organizations has considerable interest in that area, and the potential for joint funding and/or sharing of results is greatest. The Panel considered the mechanisms and strategies available to the Government to maximize its leverage and accomplish its stated goals and objectives. The panel also considered the appropriate Government responses to the "threat" that foreign programs may pose to the competitive and economic position of the United States. The FCCSET panel on research coordination plans to report periodically on Very High-Performance Computing throughout the Government and continue to meet regularly to maintain high-level coordination among Federal agencies and departments. As part of its ongoing activities report, the panel plans to assess the status and effectiveness of collaborative activities in this field.

The recommendations cited herein address key issues related to ensuring that the United States not only retains but advances its position of leadership in the Very High-Performance Computing field. Since the other FCCSET supercomputing panels have studied procurement strategies relative to the commercial market and provisions for greater access to supercomputers, these issues are not covered here.

In making its recommendations, the panel recognized that sustained and growing support for basic research is necessary for advancement in the Very High-Performance Computing field. Other related areas also need to be addressed in order to successfully move existing technology out of the laboratory and into applications. Overall, these recommendations are intended to ensure that the United States continues to be the prime source of Very High-Performance Computing technology in the decades ahead.

EXECUTIVE SUMMARY

The United States has long been recognized as the leader in the computer and information processing field and despite the challenges presented by foreign initiatives, has retained its leadership due to the strong technical, industrial, and academic base and research support that have evolved over the past decade. Continuing efforts within the research community, however, have not prevented this position from eroding rapidly. Although the United States remains dominant in information processing, the Japanese and others have targeted both high-end numerical computing, or supercomputers, and symbolic processing for technological exploitation on their quest to capture leadership in the information processing field.

Since the early 1960's, the key technical and programmatic issues in high-performance computing have been investigated by a series of studies, reports, and workshops sponsored primarily by the Federal Government. Interest in these issues declined in the mid-1970's as Government funding levels fell off, but has been revived by the perceived threat of foreign competition, the emergence of innovative architectural concepts in parallel and multiprocessor machines for scientific and symbolic computation, and the identification of a broad spectrum of mission-specific applications that will require radical improvements in the speed and performance of computing systems. Renewed Government interest and funding have spawned such activities as DARPA's Strategic Computing Program, NSF's Advanced Scientific Computing program, DOE's Energy Science Supercomputing Program, and DoD's Supercomputer Research Center. In FY85, the total Federal investment in Very High-Performance Computing research will approach $101 million.

To address the need for the next generation of both numerical and symbolic computing capabilities, the Federal program in Very High-Performance Computing encompasses a wide spectrum of activities in Advanced Computer research and development. Significant high-performance parallel and multiprocessor computing advances will come from research efforts in machine architecture, advanced programming languages, and systems and software methodologies for concurrent multiprocessor operations. Areas such as system software for parallelism--including artificial intelligence techniques applied to debugging, testing, verification, and performance measurement--will require significant investment over the next decade.

Researchers in high-performance numerical computing and symbolic processing expect significant gains in performance from advanced multiprocessor architectures. Although the design and development of multiprocessors for symbolic and numerical problems are now being explored separately for technical reasons, significant similarities in the underlying architecture concepts may emerge as basic research efforts continue. Certain numerical machines are capable of performing symbolic computations quite rapidly, but they have not been optimized for this class of applications and generally are regarded as not cost-effective. Existing high-performance computers for numerical and symbolic domains are based

widely on varying architectural concepts, but long-term trends indicate that important common, underlying aspects, with complementary domain-specific specializations, will result. It should be possible to develop supporting hardware and systems software to allow multiprocessors to operate in both symbolic and numeric environments concurrently. Further, for these systems, there is a critical need for characterizing and measuring computer performance to permit designers and users to discriminate among alternative architectures.

Federal high-performance computing research consists of a number of programs and activities carried out by individual Government agencies and by combinations of Government agencies working together. This overall effort supports the Government's mission to:

(1) Conduct basic and applied research in computational sciences and engineering, specifically in high-performance computing;

(2) Develop promising concepts into prototype systems where appropriate;

(3) Evaluate the performance of existing and planned high-performance computing technology;

(4) Ensure the necessary infrastructure for the conduct of the program; and

(5) Apply high-performance computing in meeting mission-specific goals.

Coordination among the Government agencies and departments responsible for funding advanced computer science research is carried out at the program manager level. Program managers are responsible for sharing research plans and results; reaching agreement on generic and mission-specific goals and on program interactions and modifications; discovering relationships, overlaps, gaps, and opportunities in various research activities; and creating a coherent total-program approach. Program managers throughout the Government work to get the most from their funds by sponsoring programs in a joint or complementary manner where appropriate. Although it is impossible to quantify the amount of formal or informal coordination taking place, it is substantial. Program managers routinely refer research proposals to one another if certain tasks, or the entire proposal, seem more appropriate to the other agency's missions and programs.

By providing a forum for the exchange of information about individual Government agencies' research and development programs in very high-performance numerical and symbolic computing, this panel has developed a global perspective on interagency coordination. The existing coordination process could be more visible and the panel will continue to emphasize the visibility issue as the support for very high-performance computing research and development in the Government increases.

In the course of its deliberations, the panel developed a set of recommendations to enhance the United States position as the leader in very high-performance computing. The recommendations are summarized as follows:

- o Maintain a vigorous, coordinated research program.
- o Increase emphasis on understanding fundamental issues in parallel processing.
- o Promote research activities that apply to a broad class of problems.
- o Do not over-coordinate basic research.
- o Explore a diverse set of architectures.
- o Coordinate exploratory machine architecture development efforts.
- o Improve technology transfer mechanisms from Federally sponsored research to the commercial sector.
- o Develop programs designed to augment the number of trained researchers.
- o Take steps to ensure that compensation is adequate to retain qualified researchers in the public and academic sectors.
- o Develop effective performance measurement and modeling techniques.
- o Investigate the infrastructure requirements to support the research community.
- o Maintain a visible interagency coordination effort.

The FCCSET Panel on Advanced Computer Research considered the impact of the Federal research program in advancing the United States capabilities in very high-performance computing, and the key support issues such as facilities, personnel, and Government investment in the requisite technologies. The recommendations that have been developed provide a basis for the continued development of the technical expertise the United States requires to maintain its leadership position in information processing and advanced, very high-performance computer technology.

MEMBERS OF THE FCCSET PANEL ON
ADVANCED COMPUTER RESEARCH
IN THE FEDERAL GOVERNMENT

Dr. Robert E. Kahn, Defense Advanced Research Projects Agency, Chairman

Dr. Donald Austin, Department of Energy

Mr. Ross Bainbridge, Department of Commerce, National Bureau of Standards

Dr. Bernard Chern, National Science Foundation

Dr. Ronald L. Larsen, National Aeronautics and Space Administration

Dr. Joseph Markowitz, Central Intelligence Agency

Dr. K. M. Speierman, National Security Agency

Dr. Andrew Pettifor, Office of Science and Technology Policy

Additional contributors to the report are listed in Appendix E.

REPORT OF THE FEDERAL COORDINATING COUNCIL ON SCIENCE, ENGINEERING AND TECHNOLOGY PANEL ON ADVANCED COMPUTER RESEARCH IN THE FEDERAL GOVERNMENT

A. INTRODUCTION

The United States has long been recognized as the leader in the information processing field and, despite the challenges presented by foreign initiatives, it remains in that position as a result of its strong technical industrial base and continuing research support. In spite of vigorous continuing efforts within the research community, however, this lead is being threatened. Foreign programs aimed at Very High-Performance Computing signal the need for a vigorous national response by the United States to maintain its leadership.

Since the early 1960's, the key technical and programmatic issues in Very High-Performance Computing have been documented in a series of studies, reports, and workshops sponsored primarily by the Federal Government (Lax, et. al., 1982; Feigenbaum, et al., 1983; Schwartz, et al., 1984; and Decker, et al., 1984). Renewed Government interest and funding have spawned such activities as DARPA's Strategic Computing Program, NSF's Advanced Scientific Computing Program, DOE's Energy Science Supercomputing Program, and DoD's recently announced Supercomputer Research Center. Research in Very High-Performance Computing is not yet a major segment of the overall Federal effort in Advanced Computer Science research but it is clearly growing.

The Federal Coordinating Council on Science, Engineering, and Technology Panel on Advanced Computer Research in the Federal Government was originally tasked by the Office of Science and Technology Policy to stimulate the exchange of information within the Government on symbolic computing and artificial intelligence (AI). Subsequently, the panel broadened its charter beyond symbolic processing and artificial intelligence to include research in very high-performance scientific and numerical computing, with the panel focusing on the Federal role in advancing the state of the art in this field.

The purpose of this report is to provide a concise summary of ongoing Federal research efforts in Very High-Performance Computing, and to identify a set of findings and recommendations which we believe will strengthen our nation's overall capability in this extremely important area. This panel compiled a series of interrelated findings and recommendations on the issues that Government policymakers should consider in order to ensure that the United States retains its strong national capability in advanced computing technologies. Generally, the size and diversity of the many efforts within the Government, industry, and academic community result in special problems in technology transfer, availability of qualified personnel and necessary research facilities, and

52-821 O - 86 - 22

information flow. The recommendations take into account the importance of basic research as well as a number of key engineering issues that must be considered before technologies move from the research laboratory environment into mainstream applications. The recommendations, taken as whole, will assure that the United States continues to be the prime source of Very High-Performance Computing technologies in the decades ahead.

The Panel defines Very High-Performance Computing to mean efforts specifically concerned with the exploitation of concurrency and parallel processing to achieve dramatic increases in speed of computation. The main focus is on development of multiprocessor systems with emphasis on scalable architectures whose performance increases are near linear as more processors are added. Specifically included are computational models and architectures in parallel processes; problem decomposition techniques and languages for expressing parallelism; compilers and operating systems specifically devised for multiprocessor systems and software which manages and controls them; algorithms and heuristics designed for parallel processing; applications which will lead to an increased understanding of the underlying principles of these technologies; and those methods for quantitatively characterizing and measuring advances in each of these areas.

The Government also supports a significant amount of Advanced Computer Research which is targeted at improved functionality, reliability, etc. and which the Panel has excluded from its interpretation of Very High-Performance Computing. Most computer research that leads to incremental improvement in computation speed has been excluded from this category. Most artificial intelligence research is expected to contribute major functional advances in computation, but little of it is yet addressing performance speed-up and is not included. Further, most ongoing research in basic computer science and computational mathematics, such as efforts in software technology, distributed systems, numerical analysis, and computer networking which are aimed at substantial advances in functionality, are excluded. In addition, related information processing research efforts such as software life cycle maintenance, system reliability and microelectronics technology, although vitally important in development and applications, have not been included in either category. Approximately $298 million is being spent in FY85 on Advanced Computer Research, of which approximately 34 percent or $101 million is being spent on Very High-Performance Computing research.

The Panel observed that the approaches being taken for high-performance numerical computing and for symbolic processing are noticeably different at this time, yet researchers in both areas expect significant gains in performance from multiprocessing. Existing high-performance computers for the numerical and symbolic domains vary widely, but long-term trends indicate important common architectural aspects with complementary domain-specific specializations. Certain numerical machines are capable of performing symbolic computations quite rapidly, but they have not been optimized for this class of applications and generally are not regarded as cost-effective. Although the conceptual design and machine development of multiprocessors for symbolic and numerical problems are now being explored

separately for technical reasons, significant similarities in the underlying architectures are expected to emerge later. The trend in the use of highly parallel, multiprocessor computer architectures toward combined numeric and symbolic applications may lead to architectures capable of executing both forms of computation in an efficient and cost-effective manner.

Advanced numerical computation addresses key generic areas such as:

1. Numerical analysis and simulation applied to computational models for which closed-form solutions do not exist;

2. Manipulation of very large volumes of data, perhaps generated from the numeric solution to the model analysis as in item 1; and

3. Graphić presentation of complex, multi-dimensional data.

Symbolic processing manipulates non-numeric objects and is concerned with issues such as:

1. Knowledge representation and semantic retrieval techniques;

2. Feature or symbol extraction, as in the translation of a signal into a symbolic representation;

3. Expert systems which combine knowledge with inferencing mechanisms; and

4. Search techniques based on non-exhaustive heuristic methods.

In certain cases, such as image understanding and speech recognition, numerical techniques are first used to prepare data for input to a symbolic processor, such as in generating a primal sketch, which consists of significant lines and edges in an image.

The Federal research effort in Very High-Performance Computing consists of a number of programs carried out by individual Government agencies or by combinations of Government agencies working together to meet specific national mission and application needs. Requirements for defense, space technology, energy technology, excellence in scientific research, as well as others, are reflected in the overall programmatic goals.

In the course of its deliberations, the panel focused on four key issues concerning the ability of the Government to conduct a successful basic research and exploratory development program in Very High-Performance Computing. These issues are:

1. **TECHNOLOGY** - What research areas must be exploited to advance the state-of-the-art;

2. **PERSONNEL** - How to ensure a sufficient number of researchers by the end of the decade, how will the projected shortfall be met in high-performance computing;

3. **FUNDING** - What is the current Government investment and how is it apportioned; and

4. **RELATED ISSUES** - Coordination of Federal activities, socio-economic impacts, and facilities to directly support research activities in high-performance computing.

The socio-economic environment for computer research and development in the United States is undergoing significant changes in order to meet the increasing demands of industry, Government, and the scientific community for numerical and symbolic computational capabilities. The recent formation of such entities as the Semiconductor Research Corporation (SRC), the Microelectronics and Computer Technology Corporation (MCC), the Parallel Processing Research Council, the recently announced Software Productivity Consortium (SPC), as well as the growing number of direct industry-academic partnerships in computer research and development, is in itself a benchmark of the concern being expressed by all sectors of the research community as to the strength of the United States' technical leadership position. The private investment in all categories of university research increased by 13 percent from 1982 to 1983 and exceeded $390 million. This private investment--in the form of fellowships, grants, or direct contracts--can be expected to continue to increase, especially as the programs within the joint ventures reach projected funding levels. The recently signed National Cooperative Research Act of 1984 facilitating the formation of joint research and development ventures will further promote private investment into the academic and industrial research community. On the other hand, given the charter of these joint ventures, it is not clear how the transfer of technology from the sponsored research to the commercial marketplace beyond a consortium's individual members will take place. In addition, if researchers are drawn from the academic community into projects conducted within the consortium's own staff, a decline in the ability to educate new graduate students may result as the number of qualified educators decreases.

The overall Federal program on Very High-Performance Computing research encompasses a wide spectrum of activities. Basic research and exploratory development programs in this area will lead to a new generation of very high-speed computers and provide the underlying theory and knowledge that will create an environment in which innovative ideas in information processing technology will flourish. These efforts will also be a key source of scientists and engineers required for the future growth in this rapidly expanding and critical field.

B. THE FEDERAL ROLE

The various Government agencies and organizations involved in Advanced Computer Research support both generic and mission-specific programs. This report focuses on the Federal research activities in Very High-Performance Computing with emphasis on those that are generic in nature and which yield fundamental concepts, technology, knowledge, people and ideas. Activities directed toward an individual organization's mission-specific goals have not been included in determining the amount of funding for Very High-Performance Computing research.

Research and development can be viewed as a four-part process. The first step in this process, basic research, consists of fundamental undirected exploration at the conceptual level. Research into these concepts is often initiated by the investigator doing the research utilizing existing facilities, computer resources, and personnel, funded at low cost, and accomplished with little or no special equipment; coordination on a project-by-project basis is appropriate and effective. The key to the success of basic research programs is highly dependent upon the contributions of individual researchers. Federal coordination typically takes the form of peer review within the relevant Federal agencies. A researcher who has proposed a program may even be presented with a "coordinated" Government response to his or her project, which in turn can lead to funding of selected tasks of a total proposed program by different agencies. The specific areas selected for funding reflect the overall mission specific interests of each agency. However, this form of coordination encourages the entrepreneurial aspects of the research community to propose new areas for exploration, even if they go beyond the individual interests of the single agency or department of the Government.

The second part of the process, exploratory development leading to experimental capabilities, is concerned with the feasibility of applying promising basic research results, often to generic classes of problems. In the hardware domain, a prototype or "breadboard" might be constructed along with elementary system software. At this point research begins to require increasing capital resources and facilities; the need for more extensive coordination begins to emerge. Multiple parallel efforts, although often desirable at the basic research level, may be unaffordable for exploratory development, so refinement and selectivity may be necessary here. On occasion, several sponsors may pool their resources toward a common goal, or each may support complementary aspects of a program. Examples include the Cosmic Cube (architecture funded by DARPA, applications by DOE), the Wave-front Array Processor (basic research funded by NSF, signal processing application by ONR), and Systolic Array Processors (computational model and algorithm research funded by NSF and ONR, architecture and breadboard by DARPA).

The third and fourth phases of the process involve advanced development and production engineering. At these levels, program costs are far greater and only a few efforts can typically be pursued. The selection of

candidates for advanced development is done individually by each organization in response to mission-specific needs. The need for coordination across organizational lines rapidly decreases as the technologies move into these latter phases.

The Federal Program structure in Advanced Computer Research is depicted in Figure 1 as a series of building blocks or supporting layers of which the research program in Very High-Performance Computing is a part. As indicated above, emphasis was placed on programs that are generic in nature. Mission-specific programs were not included in determining the overall Federal funding profile. Also, programs related directly to purchasing computers or facilities for <u>other</u> scientific and engineering disciplines were not included in any of the research figures, an example being the recently announced NSF program on advanced scientific computing.

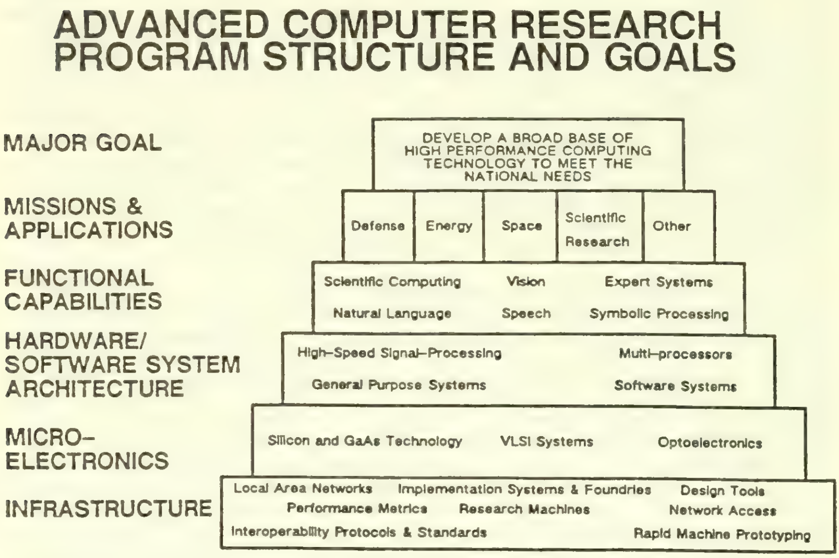


Figure 1

Federal agencies interact in a variety of ways to coordinate their efforts, including:

1. Seeking agreement on generic and mission-specific goals;

2. Sharing research plans and results;

3. Discovering relationships, potentially overlapping interests, and opportunities in various research projects;

4. Conducting joint research programs; and

5. Creating a coherent, combined research program encompassing Government, industry, and academic sectors while maintaining mission-specific goals of individual organizations.

For many years, coordination among the agencies and departments responsible for funding computer science research has been effectively carried out at the program manager level. More recently, the FCCSET panels have provided a higher level of coordination as well. Although it is impossible to quantify the existing formal, semiformal, and informal interagency coordination, and although it is not completely visible outside the participating organizations, the amount of coordination is substantial. The authors of this report recognize the importance of providing increased visibility to this process as the support for Very High-Performance Computing grows. The report of this panel is expected to assist the coordination process by summarizing the ongoing and planned high-performance computing research and development activities in the Government. The panel itself is the most visible example of inter-agency coordination and visibility into the coordination process will be enhanced by the regular reviews of the FCCSET panel.

C. DEFINITION OF CATEGORIES

Advanced computer research is comprised of a set of multidisciplinary basic research and exploratory development activities, for which the panel developed the following set of nine categories to characterize them:

1. **Computational Mathematics:** The design, analysis, and implementation of algorithms for solving basic numerical problems by computer.

2. **Computer Architecture (Hardware and Software):** The design, simulation, and development of new computer architectures including both hardware and system software.

3. **Machine Intelligence and Robotics:** The development of software and conceptual designs which allow computers to carry out tasks which would be considered "intelligent" if performed by a human.

4. **Distributed Computing and Software Systems:** Techniques and procedures for building systems consisting of multiple computers connected by communication networks; the technology for designing and building software systems, including rapid prototyping and reliable operation.

5. **VLSI Design and Special Purpose Computing:** Tools and techniques for designing state-of-the-art VLSI, and the development of innovative circuits and computing systems using those tools and techniques.

6. **Data Management:** Design and development of advanced data base management concepts and systems.

7. **Theoretical Computer Science:** The analytic study of fundamental problems in Computer Science.

8. **Network and Research Facilities:** The provision of facilities for communication and computation specifically for the purpose of furthering advanced computer science research.

9. **Performance Evaluation and Modeling:** The analysis and study of performance evaluation, metrics and models, operational standards and benchmarks, user interface technologies, and human factors.

These research areas are not entirely separate and distinct but interdependent. For example, studies in theoretical computer science may lead to the understanding of techniques to be applied by computational mathematics, which in turn can lead to new approaches or tools for constructing highly complex VLSI designs for multiprocessor architectures. Network and research facilities generally support the work in the other categories.

Figure 2 shows Very High-Performance Computing depicted as a small part of Advanced Computer Research which, in turn, is a small part of Information Processing research and development. The definition of Advanced Computer Research is itself subjective. The shaded area in Figure 2 indicates the panel's qualitative assessment of the coverage of the research activities, and thus the funding profiles, in this report under Advanced Computer Research. Programs may span across the boundary between the regions. In identifying programs and categorizing the relevant research activities, the result may omit some efforts not totally within Advanced Computer Research or may include those elements of Advanced Computer Research programs that are outside of the Advanced Computer Research arena. For example, the DoD STARS and ADA programs were judged part of the larger information processing category, as was much of the computer network research. Program application or mission specific computer research activities were specifically not included in this report. In addition, none of the proposed SDI efforts were included since at the time this report was prepared, that program was still being defined. In developing the funding profiles for Advanced Computer Research and Very High-Performance Computing Research, we were unable to

insure systematically that all such funds were properly categorized. We believe the funding data presented in this report represent quite close estimates, based upon the information available to the Panel. In most cases, research efforts were decomposed into several tasks in order to allocate the funding by the categories identified by the Panel. As such, the specific funding breakdown presented may not reflect actual categories of funding budgeted by the various organizations.

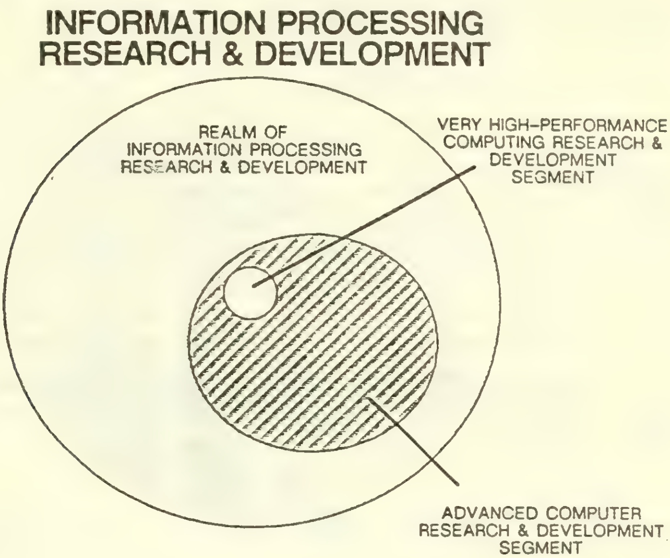


Figure 2

The panel compiled funding numbers for both Advanced Computer Research and for Very High-Performance Computing but focused its coordination efforts on the latter. Parallel computing architectures, many of which are still highly experimental, promise large increases in computational power without significant redesign of their components--that is, they are "scalable." Such architectures may make possible machines that are several orders of magnitude faster than the fastest existing machines. Scalable parallel architectures have major potential advantages in a VLSI environment because they can be built from large numbers of identical parts that can be mass-produced efficiently, because the design cost can be amortized over multiple configurations differing only in size, and because they are cost-effective.

The development of many small scale parallel experimental machines is seen as a necessity to explore all the promising ideas in multiprocessor architectures. Although device speeds continue to increase as the minimum feature size decreases, the physical limits of MOS technology will soon be reached. Multiprocessors offer the possibility of dramatic increases in speed through parallelism. There is a critical need for characterizing, measuring and modeling computer performance to permit designers and users to discriminate among alternative architectures. The speed of technology development in industry is increasing, and we are beginning to witness a new class of mini-supercomputers emerge in the marketplace. By the end of the decade, we expect to see the power of today's fastest supercomputers at the price of today's most powerful minicomputers.

D. FEDERAL ACTIVITIES AND FUNDING

The Federal investment for FY 1983, 1984, and 1985 in Advanced Computer and Very High-Performance Computing Research is shown in Figures 3 and 4, respectively, by the major funding organizations. Individual organization funding is summarized in Appendix A, Tables A-1 and A-2. To further delineate the Federal investment in these two areas, the panel reviewed the Federal efforts in each of the nine categories in Section C. The expenditures in these research categories in FY 1983, FY 1984, and FY 1985 are shown for Advanced Computer and Very High-Performance Computing Research in Figures 5 and 6, respecitvely, with detailed funding by area given in Appendix A, Table A-3. The total funding by organization and research category was compared, and the results are summarized below.

It has been difficult to develop precise funding figures for both Advanced Computer Research and Very High-Performance Computing since most Federal programs are not fiscally structured along those lines. The figures listed herein reflect the best estimates of spending in each of the categories.

The Federal investment in Very High-Performance Computing Research was approximately \$36.6 million in FY 1983, \$57.8 million in FY 1984 and \$100.9 million FY 1985 as shown in Table A-2. Figure 7 graphically shows the relative Federal investment in the Very High-Performance Computing component as a part of the overall Advanced Computing Research activity. Very High-Performance Computing Research comprised 21.1, 25.5, and 33.8 percent of the Advanced Computing Research funding in FY 1983, FY 1984, and FY 1985, respectively. The total growth rate for Very High-Performance Computing Research funding has exceeded by approximately a factor of 1.6 the growth rate in the overall Advanced Computer Research program from FY 1983 to FY 1985.

The university community forms a significant portion of the resources that are critical to the success of the Federal program in Very High-Performance Computing Research. Tables A-1 and A-2 present, for comparison, the Federal funding in Advanced and Very High-Performance

FEDERAL INVESTMENT IN ADVANCED COMPUTER RESEARCH BY ORGANIZATION, FY 1983–1985

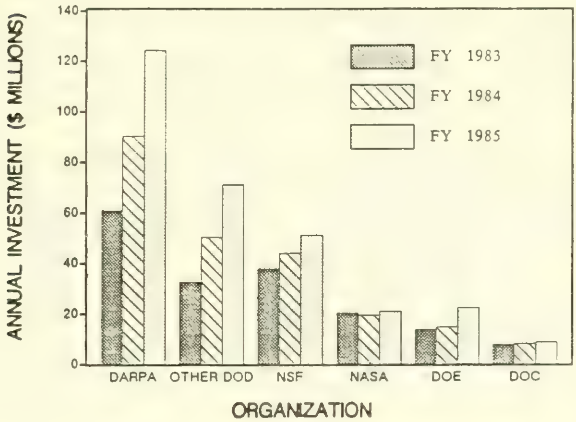


Figure 3

FEDERAL INVESTMENT IN VERY HIGH-PERFORMANCE COMPUTER RESEARCH BY ORGANIZATION, FY 1983–1985

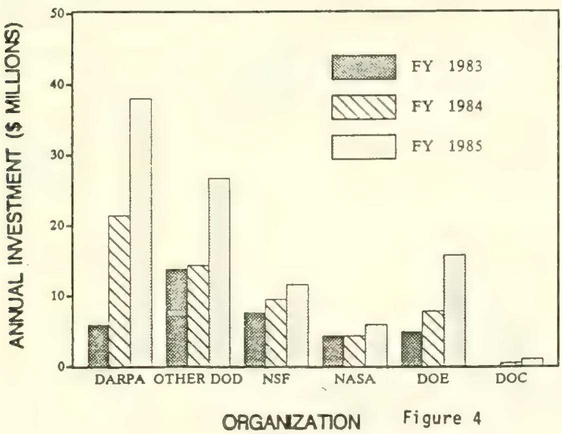


Figure 4

FEDERAL INVESTMENT IN ADVANCED COMPUTER RESEARCH BY RESEARCH AREA, FY 1983–1985

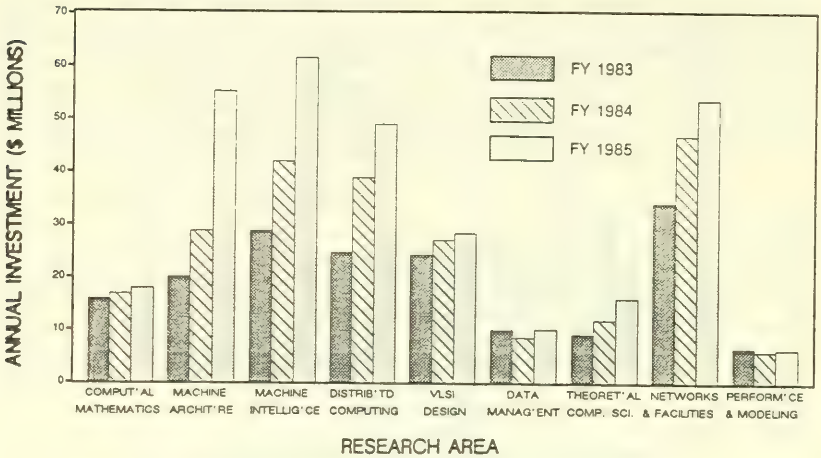


Figure 5

FEDERAL INVESTMENT IN VERY HIGH-PERFORMANCE COMPUTING RESEARCH BY RESEARCH AREA, FY 1983–1985

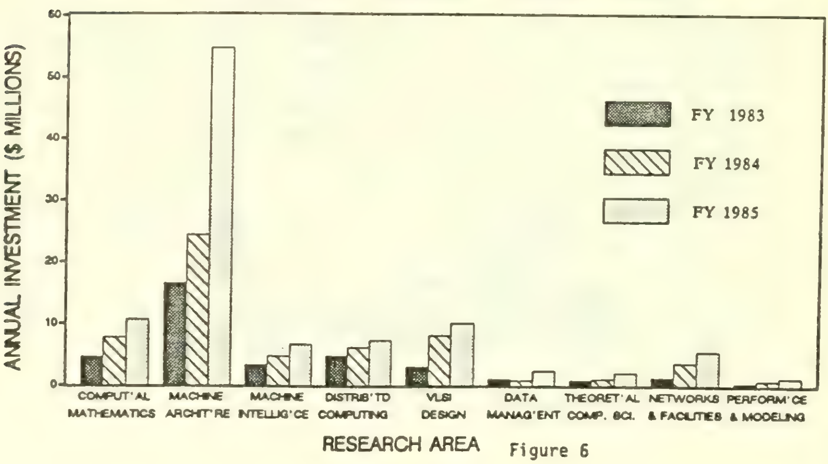


Figure 6

Computing Research, respectively, that is directed to the university sector. In FY 1985, it is projected that university funding will total $180.0 million, comprising approximately 60.3 percent of the budget for Advanced Computer Research, and approach $55.9 million, or 55.4 percent, of the Very High-Performance Computing Research budget. The university community has seen an increase from FY 1983 to FY 1985 of $67.5 million in its annual funding in the overall Advanced Computer Research program and an increase of $30.8 million in its annual budget in Very High-Performance Computing Research. Figure 7 also presents the relationship of the university segment to the overall funding profiles in this field.

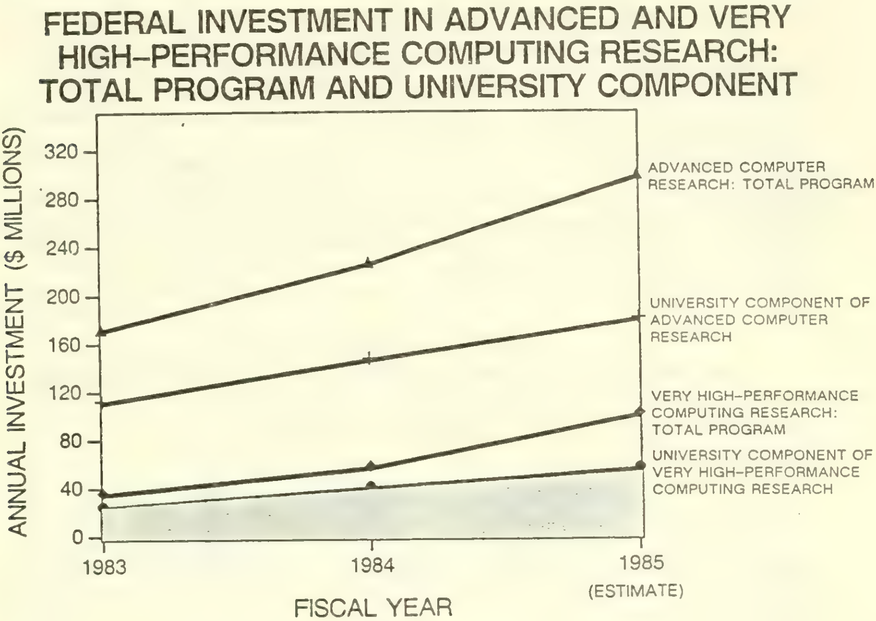


Figure 7

The funding of research in the area of machine architectures has exhibited the largest absolute growth of any of the nine categories. This growth is a reflection of the importance and the challenge that now confronts present day machines to meet the information processing requirements demanded by both civil and military systems. Specifically, in the multiprocessor area, several dozen studies and small scale explorations are underway. The Schwartz report (J. T. Schwartz, et al.,

"Report of the Research Briefing Panel on Computer Architecture," 1984) noted the tradeoff between the number of processors comprising a computer architecture and their size and complexity (or "granularity"). A coarse-grain design uses a relatively small number of high-speed processors, typically a microprocessor available as a commercial product. A fine-grain design strives for as much parallelism as possible with a large number of small, but carefully designed elementary processors each of which is fast but low in overall complexity. In general, these machines can execute multiple instructions simultaneously depending on their architecture. They may have a single memory address space shared by all processors or a separate memory address space per processor.

In FY 1984, Federal agencies were conducting research into thirty coarse-grain and eight fine-grain architectures. The performing organization and the Federal sponsoring agencies are listed in Appendix D, Tables D-1 and D-2 for coarse-grain and fine-grain architectures, respectively. A summary of the Federal investment in coarse- and fine-grain multiprocessor architecture research is shown in Table 1. The investment totals for fine- and coarse-grain architecture programs shown in Table 1 represent a significant portion of the total investment in machine architecture. In general, the fine- and coarse-grain totals reflect experimental prototype machines, simulation and exploratory development, whereas the remainder of the funding as shown in Table A-3 addresses conceptual architecture research.

TABLE 1

FEDERAL INVESTMENT IN ADVANCED MULTIPROCESSOR ARCHITECTURE DESIGNS

(FY 1984, $ MILLIONS)

	COARSE-GRAIN DESIGNS		FINE-GRAIN DESIGNS	
AGENCY	NUMBER OF PROGRAMS	FUNDING	NUMBER OF PROGRAMS	FUNDING
DARPA	5	$5.3	4	$4.5
AIR FORCE	2	0.4		
NAVY	4	3.5	1	0.4
NSF	17	4.4	3	0.4
DOE	9	2.6		
NASA	7	0.9	1	1.0
TOTAL	30*	$17.1	8*	$6.3

* Not may not equal column total due to joint programs.

The Federal funding data for the fiscal years 1983, 1984, and 1985 for Advanced Computer Research by organization is detailed in Appendix B, while Appendix C presents similar data for Very High-Performance Computing Research, in accordance with the nine research categories described in Section C. From these tables, areas of significance to each agency can be seen, along with the trends in each research category.

The increase in Government funding between FY 1983 and FY 1985 has been both in existing programs and in funding of new technology initiatives. The Strategic Computing Program, the Supercomputing Research Center, the Computer Measurement Research Facility, and the Energy Science Advanced Computation Programs are all new initiatives within the past two years and are funded at significant levels.

A wide range of jointly funded and coordinated projects to support research and development in high-performance computing is already in progress. A summary of the several major efforts follows.

1. Networking

Various network projects have been carried out jointly, or the results have been shared. The ARPANET/MILNET is used by DoD, DOE, NASA, NBS, NSF, and other parts of the Government. The CSNET activity of NSF is a joint effort supporting the entire computer science research community. The various networking activities continue to be closely coordinated as new networks become available.

2. Foundry Services

A joint effort has been undertaken by the National Science Foundation and the Defense Advanced Research Projects Agency that allows U.S. universities to use the DARPA-developed MOSIS fast turnaround VLSI implementation facility for university-based research and educational programs requiring the fabrication of digital designs as integrated circuit chips. Designs submitted by DARPA or NSF researchers in digital form over the ARPANET or Telenet using standard design rules and a standard artwork format are fabricated and returned to designers in four to six weeks. Both NMOS and CMOS are widely available, along with printed circuit board services. With appropriate authorization, it is also possible to provide access to VHSIC technology. The fabrication service is well documented, and other Government-sponsored programs are being encouraged to use the system.

3. Resources

CRAY machines at several locations are being made available for shared use. It is expected that the NSF supercomputing sites will be available for shared usage and quid pro quo arrangements should be explored for access to other systems. In addition, plans are in progress

to provide access, including ARPANET and CSNET systems, to new numerical and symbolic processors as prototype machines become available from the many Federal projects in machine architecture and machine intelligence.

4. Architecture

1. The 128-node Butterfly multiprocessor built by Bolt, Beranek, and Newman (BBN) for DARPA was selected by the University of Rochester for experimental research funded by NSF.

2. The Pixel Planes Graphics projects at the University of North Carolina was jointly funded by NSF and DARPA.

3. The 4.2 BSD Unix software developed by Berkeley for DARPA has been distributed to most major U.S. universities for their research.

4. Data flow research at MIT and the Cosmic Cube at California Institute of Technology is jointly supported by DOE and DARPA.

5. Of the 38 architecture programs listed in Tables 1 and 2, 13 (or 34 percent) are funded by multiple agencies.

E. PERSONNEL

Additional highly qualified research scientists and engineers will be needed to sustain the desired growth rate in high-performance computing. Growth in university faculty is essential to ensure an increasing pool of quality graduate students. Further, the number of qualified technical personnel for research in university, industry, and Government laboratories, as well as those with leadership capabilities to act as program managers within the Government, must be increased. The management and direction, as well as the actual research efforts, of basic research are set by persons who must be technically qualified to assess the accuracy and the quality of a proposed research program. This is particularly true of the program management personnel in the various Federal agencies involved in high-performance computing research. However, the number of new Ph.D.'s in computer science and information processing has stayed relatively constant since 1976. The increasing number of commercial successes in the field will create attractive opportunities to drain away both scientific personnel and funding from basic research areas which could lead to an overall decline in the rate of progress in the field and divert new graduates away from academic careers.

It is too soon to tell whether the number of Ph.D.'s will increase as a result of the increasing number of Bachelor's and Master's degrees being granted, since students from those classes who may be seeking Ph.D.'s have not yet completed their study. However, given earlier statistics, it would appear that without some method of encouragement, the number of

Ph.D.'s in computer science and related fields will not increase significantly in the future. If advanced research in Very High-Performance Computing is to continue and expand, these numbers must be increased. Without question, many of those persons with Bachelor's degrees and the ability to continue further in their education choose to go into industry instead. Indeed, a particularly disturbing trend is that some colleges and universities are considering limiting class size in undergraduate computer science and electrical engineering since the demands in these areas are stretching faculty resources and leaving less time for research. It is not possible to predict with accuracy the consequences of such actions on the role of development in high-performance computing.

Universities have increased their salaries substantially in the last several years to make the university environment more attractive. Starting salaries at many institutions are now comparable to those of industry; however, there is a problem in salary compression at the higher ranks in universities. Salaries in the public sector are lagging behind the levels in both academia and industry. Retaining high-quality personnel in universities and Government in the face of strong industrial incentives presents a particularly important challenge.

Industry lures not only newly graduated scientists and engineers, but also members of the existing, limited pool of technically qualified researchers in Very High-Performance Computing Research. The impetus for these people to move to the private sector to capitalize upon their research work can easily be seen in the rapidly increasing availability of venture capital to initiate new start-up companies.

The problem within the Government has been acknowledged to a limited extent. One attempt to provide an incentive for individuals to join or remain with the Government, at lower pay-scale levels, has been made by the use of a pay differential for critical engineering disciplines. Also annual increases have been given early in certain critical specialty areas. While recent improvements in Government pay for engineers have been beneficial, we note that computer scientists, broadly described, comprise an even scarcer category of skills. Whether a computer scientist has an engineering degree depends upon cultural factors at particular universities and not factors intrinsic to the discipline or the required education. The definition of those who qualify for engineering pay differentials should be broadened if the Government is to compete with industry in attracting qualified computer scientists.

Access to high-performance computing facilities by a wide range of scientists and engineers is considered an absolute necessity for areas in which major advances are now completely dependent on access to the most advanced scientific computational resources. Dedicated supercomputer systems exist at a small number of Government laboratories and research centers. Many of these systems are already saturated, moreover, they serve only a small portion of the nation's research community. NSF has estimated that the demand for access to supercomputing resources by the

research community now exceeds the capacity by a factor of three. The overall issue of access and procurement of high-performance computing facilities is addressed more fully by the FCCSET Panel on Supercomputer Procurement and Access chaired by DOE.

F. FINDINGS AND RECOMMENDATIONS

This panel compiled a series of interrelated findings and recommendations on the issues that should be considered in order to ensure that the United States retains its strong national capability in advanced computing technologies. Other FCCSET panels have studied procurement and access strategies to drive the commercial supercomputer market and make supercomputer resources available to researchers and thus these issues are not repeated here.

The following recommendations take into account the importance of basic research as well as a number of significant engineering issues that must be considered immediately to speed existing technologies from the research laboratory into the mainstream of applications. The recommendations will assure that the United States continues to be the prime source of computing technologies in the decades ahead.

The following are the findings determined by the panel during the course of the discussions:

1. Federal support for Advanced Computer Research and Very High-Performance Computing has grown significantly in recent years and a vigorous research and development program is developing. The Federal program in Very High-Performance Computing is a collection of individual efforts by various departments and agencies. This effort is critical to provide the advanced computing capabilities that will be needed in the future. Despite the growth in the past two years in the level of Federal funding and the number of individual programs being conducted in this field, there are still significant areas that are not funded but appear to have promise for exploitation.

2. Despite significant engineering expertise in multiprocessors, and computer technology in general, there is a lack of underlying theoretical understanding of Very High-Performance Computing. Efforts need to be expanded in understanding such areas as multiprocessor architectures, languages, software systems, algorithms, problem decomposition techniques, etc., as well as in the applications that will be executed on new multiprocessor based computers. Equally large gains in performance can come from attention to software and algorithms as from hardware speed and the performance of specific architectures. The Federal investment in this critical area has been insufficient.

3. The agencies which sponsor research in Advanced and Very High-Performance Computing necessarily give priority to their own needs and mission-specific goals (e.g., reliability, hardening, speed, and power). Basic research proposals are typically formulated by individual researchers and presented to various agencies. Although a typical effort may be funded by a single agency, the possibility of funding by multiple agencies allows for a greater latitude in options to finance a complete, coherent program. The optimum result of multiple-agency investment is technology or concepts that are applicable to a broad spectrum of mission-specific goals.

4. Technical program managers coordinate research on a case-by-case basis, and the method is effective. Within the Government, program managers keep each other well informed about their individual programs and goals. Regular exchange of information occurs among the agencies of the Government concerned with the specific research area, so that local optimization of specific programs occurs.

5. Multiprocessor based computer architectures appear to be so fundamental to success in Very High-Performance Computing that a thorough investigation of all promising architectures in this area is warranted. A wide variety of promising architectures have been identified. It is particularly desirable to develop machines with the intrinsic capability to provide both numeric and symbolic processing.

6. As computer research programs mature, access to advanced computing facilities, hardware, and knowledgeable scientists and engineers becomes critical to the success of the program. Given the limited availability of these key resources in the areas of Advanced Computer Research and Very High-Performance Computing, the level of coordination required must be increased to ensure the optimum allocation of these resources to a number of individual programs.

7. Although Federally sponsored research is intrinsically useful to the Government by definition, industry often lacks the economic motivation to capitalize on the results of this research. As a result, domestic research may be better used outside the United States than within it.

8. There are too few newly qualified research scientists and engineers in the United States to sustain the desired growth rate in Advanced and Very High-Performance Computing research through the end of the decade. Over the past decade, the number of graduating Ph.D.'s in computer science and engineering, as well as in the overall information processing field, has remained relatively constant.

9. To make the academic environment more attractive, universities have increased their salaries substantially in the past two years. At many institutions, starting salaries are now comparable with those of industry. Within the Government, as well as in the university sector, the problem of salary compression at the higher ranks exists. Further, Government salaries have lagged behind those in the academic and industrial sectors.

10. The ability to evaluate the performance of multiprocessor architectures is the key to sound, long-term research in this area, and may provide a solid foundation for their application. Current performance evaluation and modeling techniques, as well as standardized benchmarks, are inadequate for the emerging multiprocessor architectures.

11. The underlying support structure, or infrastructure, for research in Advanced and Very High-Performance computer research forms a key aspect of the overall program in this area. The availability of research facilities, access to remote computer facilities, rapid prototyping of computer architectures and supporting electronics and mechanical structures, and the development of common design tools, performance benchmarks, interoperability standards, and communication protocols, etc., are essential for the timely development of advanced computer structures and information processing technologies.

12. Although the coordination of basic research programs has been successfully conducted at the program manager level, it is difficult to quantify the existing formal, semiformal, and ad-hoc interagency coordination. Coordination efforts are not clearly visible outside of the participating organizations. Visibility into the coordinating activities for this critical area of research must be enhanced at the interagency level. As a National response is formulated to the perceived threat to United States leadership in information processing technology, the ability to effectively present Government activities in Advanced Computer Research is essential.

To address the issues raised in the panel's findings, the following set of recommendations have been developed:

1. **MAINTAIN A VIGOROUS, COORDINATED RESEARCH PROGRAM.**

   The United States should maintain a coordinated program of research and development in Advanced Computer Research and Very High-Performance Computing, in particular, to ensure our national defense, to foster scientific excellence, and to enhance economic competitiveness. This program is growing and should continue to do so.

2. INCREASE EMPHASIS ON UNDERSTANDING FUNDAMENTAL ISSUES IN PARALLEL PROCESSING.

   If industry and Government are to be able to fully apply new capabilities in Very High-Performance Computing, increased emphasis must be placed on achieving a better understanding of the fundamental issues in parallel processing such as algorithm development, problem decomposition techniques, languages, operating systems, and interprocess communication and synchronization mechanisms.

3. PROMOTE RESEARCH ACTIVITIES THAT APPLY TO A BROAD CLASS OF PROBLEMS.

   The Government should adopt a policy that ensures that Federally supported research is as generic as possible, commensurate with meeting unique Government needs. Federal support in Very High-Performance Computing should be strengthened through the addition of generic research initiatives above and beyond those that are mission-specific, particularly in support of the research infrastructure.

4. DO NOT OVER-COORDINATE BASIC RESEARCH.

   The Panel further recommends that no effort be undertaken at this time to force a more global optimization of the Federal basic research expenditures in Very High-Performance Computing because the field is in its infancy and moving ahead very rapidly. It is the Federal Government's role to fund long-term, high-risk basic research. Multiplicity of funding sources in basic research has been one of the important mechanisms to assure promising research is supported. This is particularly important in rapidly changing fields.

5. EXPLORE A DIVERSE SET OF ARCHITECTURES.

   A Federal investment strategy must be followed which will allow exploration of the many promising diverse parallel processing architectures at both the exploratory and advanced development stages. A great deal will be learned by investigating a large variety of real problems and applications on these new architectures. Special effort should be undertaken to identify underlying structures that may be common to both scientific and symbolic processing.

6. COORDINATE EXPLORATORY MACHINE ARCHITECTURE DEVELOPMENT EFFORTS.

   Programs to develop novel, exploratory machines should have increased coordination because significant resources are required and the leverage from interagency cooperation is greatest. This is the largest absolute growth area in Advanced Computer Research, having increased from $20 million in 1983 to an estimated $55.2 million in 1985.

7. IMPROVE TECHNOLOGY TRANSFER MECHANISMS FROM FEDERALLY SPONSORED RESEARCH TO THE COMMERCIAL SECTOR.

   Technology transfer mechanisms need to be developed to aid both industry and Government in applying the results of research and development in Very High-Performance Computing. The FCCSET Panel should assess the difficulties and successes in applying this research to mission-specific areas and commercial applications.

8. DEVELOP PROGRAMS DESIGNED TO AUGMENT THE NUMBER OF TRAINED RESEARCHERS.

   The Government should develop explicit programs aimed at training students in the area of Very High-Performance Computing, and computer science and engineering more generally. Consideration should be given to graduate-level programs, and feeder programs, with the goal of at least quadrupling the number of graduates with advanced degrees within a decade. Each agency should be encouraged to provide funds for grants to young researchers to increase the number of potential leaders in the field. The Presidential Young Investigator (PYI) awards might be used as a model for postgraduate-level candidates, but efforts to stimulate leadership at all levels of the educational process should be strongly encouraged.

9. TAKE STEPS TO ENSURE THAT COMPENSATION IS ADEQUATE TO RETAIN QUALIFIED RESEARCHERS IN THE PUBLIC AND ACADEMIC SECTORS.

   Congressional support should be sought for a program that seeks to compensate highly qualified technical personnel who are attracted to Government service but would otherwise be unavailable because of existing salary differentials between the Government and private sectors.

10. DEVELOP EFFECTIVE PERFORMANCE MEASUREMENT AND MODELING TECHNIQUES.

    A high priority should be established for increased research on computer performance metrics and modeling, and standards for benchmarking. The possibility of international cooperation in the development of common standards should be investigated.

11. **INVESTIGATE THE INFRASTRUCTURE REQUIREMENTS TO SUPPORT THE RESEARCH COMMUNITY.**

    There are a number of services and facilities that are commonly required in advanced research and development programs. The Panel recommends that the infrastructure requirements be explored with a goal of developing a specific investment strategy in this area, with support from both the public and private sectors.

12. **MAINTAIN A VISIBLE INTERAGENCY COORDINATION EFFORT.**

    We believe that interagency coordination of Federal Very High-Performance Computing Research and development activities is valuable and should be continued. This FCCSET Panel could continue to function for this purpose.

G. CONCLUDING REMARKS

The technological performance of the United States economy is fully dependent upon the industrial sector. The orientation of private research and development funds is critical to bring new technologies to the marketplace. The United States has been able to retain its leadership in high-performance computing and information processing due to the strong domestic industrial base. A major element in the Government's role is in creating the environment to foster the innovations necessary to maintain this industrial base.

Policies, when translated into specific programs, are the mechanisms to create the required environment, especially for basic and applied research, or exploratory development. Government policies fall into five areas influencing innovation (reference, "Federal Support for R&D Innovation, CBO Study, April 1984), such as:

1. **Microeconomics** - Fostering positive economic growth;

2. **Competition** - Minimizing national and international disincentives for innovation, such as trade barriers to protect U.S. industry;

3. **Tax Incentives** - Policies to allow recovery of R&D expenses;

4. **Regulatory Policies** - Influence of policies that may divert corporate funds away from R&D programs; and

5. **Institutional and Informational Support** - Enhancement of technology transfer from Federally sponsored research and development to the private sector.

Government policies in each of these areas can have either positive or negative effects upon innovation. Effective policies will allow the Government to leverage its portion of the research and development funding to meet its long-term mission-specific goals.

This Panel has brought together the primary Federal organizations funding basic research and exploratory development in Advanced Computer Research to determine the Federal activities in Advanced Computer Research and Very High-Performance Computing. This has required an in-depth understanding of the broad spectrum of technologies involved as well as knowledge of the mission-specific goals of the organizations which sponsor this research. In reviewing the investment profiles for the various agencies sponsoring work in this field, it is obvious that although the funding has been substantial and is increasing, there are still significant areas of research that are not being explored. The recommendations that are presented in this report form a basis for a sustained and vigorous program of research and development that is required to maintain the preeminent position of the United States in Advanced Computer Research and information processing technology.

APPENDIX A

SUMMARY OF FEDERAL INVESTMENT IN BASIC RESEARCH AND EXPLORATORY DEVELOPMENT IN ADVANCED COMPUTER AND VERY HIGH-PERFORMANCE COMPUTING RESEARCH

TABLE A-1

FEDERAL INVESTMENT IN BASIC RESEARCH
AND EXPLORATORY DEVELOPMENT FOR
ADVANCED COMPUTER RESEARCH
(BY ORGANIZATION/DIVISIONS)
(IN MILLIONS, CURRENT YEAR DOLLARS)

	TOTAL PROGRAM			UNIVERSITY COMPONENT		
	FY83 ACTUAL	FY84 ACTUAL	FY85 ESTIMATED	FY83 ACTUAL	FY84 ACTUAL	FY85 ESTIMATED
DOD						
DARPA	60.9	90.1	124.1	44.8	67.4	86.2
ARMY						
CECOM	2.5	6.0	6.0	0.2	0.2	0.3
ARO	2.6	6.0	6.9	2.5	5.7	6.5
TOTAL ARMY	5.1	12.0	12.9	2.7	5.9	6.8
AIR FORCE						
AFOSR	6.5	7.3	8.2	5.2	5.8	6.6
RADC	8.1	13.1	12.6	1.2	2.0	1.9
AFAL	1.4	1.7	1.8	0.4	0.5	0.5
TOTAL AIR FORCE	16.0	22.1	22.6	6.8	8.3	9.0
NAVY						
ONR	6.6	7.2	7.6	6.3	6.8	7.2
NAVELEX	5.0	6.1	6.7	0.4	0.5	0.6
NAVMAT	-0-	3.0	9.2	-0-	-0-	-0-
TOTAL NAVY	11.6	16.3	23.5	6.7	7.3	7.8
SUPERCOMPUTING RESEARCH CENTER	-0-	-0-	12.0	-0-	-0-	-0-
TOTAL DOD	93.6	140.5	195.1	61.0	88.9	109.8
NSF	37.8	44.0	51.1	37.8	44.0	51.1
NASA	20.3	19.5	20.9	8.1	7.8	8.4
DOC	7.8	8.2	8.8	-0-	-0-	-0-
DOE	13.9	14.7	22.4	5.6	6.6	10.7
TOTAL	173.4	226.9	298.3	112.5	147.3	180.0

TABLE A-2

FEDERAL INVESTMENT IN BASIC RESEARCH
AND EXPLORATORY DEVELOPMENT FOR
VERY HIGH-PERFORMANCE COMPUTER RESEARCH
(BY ORGANIZATION/DIVISIONS)
(IN MILLIONS, CURRENT YEAR DOLLARS)

	TOTAL PROGRAM			UNIVERSITY COMPONENT		
	FY83 ACTUAL	FY84 ACTUAL	FY85 ESTIMATED	FY83 ACTUAL	FY84 ACTUAL	FY85 ESTIMATED
DOD						
DARPA	5.9	21.4	38.0	4.2	15.0	21.9
ARMY:						
CECOM	-0-	-0-	-0-	-0-	-0-	-0-
ARO	1.2	1.8	2.2	1.1	1.7	2.1
TOTAL ARMY:	1.2	1.8	2.2	1.1	1.7	2.1
AIR FORCE:						
AFOSR	2.5	2.8	3.3	2.0	2.2	2.6
RADC	-0-	-0-	-0-	-0-	-0-	-0-
AFAL	1.4	1.7	1.8	0.4	0.5	0.5
TOTAL AIR FORCE:	3.9	4.5	5.1	2.4	2.7	3.1
NAVY:						
ONR	6.6	7.2	7.6	6.3	6.8	7.2
NAVELEX	2.1	0.8	1.8	-0-	-0-	-0-
NAVMAT	-0-	-0-	-0-	-0-	-0-	-0-
TOTAL NAVY:	8.7	8.0	9.4	6.3	6.8	7.2
SUPERCOMPUTING RESEARCH CENTER	-0-	-0-	12.0	-0-	-0-	-0-
TOTAL DOD	19.7	35.7	66.7	14.0	26.2	34.3
NSF	7.6	9.5	11.6	7.6	9.5	11.6
NASA	4.3	4.3	5.9	1.5	1.5	2.1
DOC	0.1	0.5	1.0	-0-	-0-	-0-
DOE	4.9	7.8	15.7	2.0	3.5	7.9
TOTAL	36.6	57.8	100.9	25.1	40.7	55.9

TABLE A-3

RESEARCH AREA FUNDING SUMMARY FOR ADVANCED COMPUTER AND VERY HIGH-PERFORMANCE COMPUTING RESEARCH

(IN MILLIONS, CURRENT YEAR DOLLARS)

	ADVANCED COMPUTER RESEARCH			VERY HIGH-PERFORMANCE COMPUTING COMPONENT OF ADVANCED COMPUTING		
RESEARCH AREA	FY83 ACTUAL	FY84 ACTUAL	FY85 ESTIMATED	FY83 ACTUAL	FY84 ACTUAL	FY85 ESTIMATED
Computational Mathematics	15.8	16.8	17.9	4.7	7.8	10.7
Machine Architecture	20.0	28.8	55.2	16.8	24.4	54.8
Machine Intelligence and Robotics	28.8	42.0	61.6	3.4	4.8	6.7
Distributed Computing and Software Systems	24.7	38.9	49.0	4.9	6.2	7.3
VLSI Design and Special Purpose Computing	24.2	27.0	28.3	3.1	8.1	10.1
Data Management	10.0	8.6	10.2	1.1	0.9	2.4
Theoretical Computer Science	9.2	11.9	16.0	0.9	1.1	2.1
Network and Research Facilities	34.1	46.9	53.7	1.3	3.7	5.5
Performance Evaluation and Modeling	6.6	6.0	6.4	0.4	0.8	1.3
TOTAL PROGRAM	173.4	226.9	298.3	36.6	57.8	100.9
VHPC PERCENT OF ACR:				21.1	25.5	33.8

APPENDIX B

ADVANCED COMPUTING RESEARCH AREA FUNDING
SUMMARY BY ORGANIZATION
FISCAL YEARS 1983, 1984, and 1985

TABLE B-1

ADVANCED COMPUTING RESEARCH AREA FUNDING SUMMARY BY ORGANIZATION

FY 1983 - ACTUAL
(CURRENT DOLLARS, IN MILLIONS)

	COMPUTATIONAL MATHEMATICS	MACHINE ARCHITECTURE (H/W & S/W)	MACHINE INTELLIGENCE AND ROBOTICS	DISTRIBUTED COMPUTING AND SOFTWARE SYSTEMS	VLSI DESIGN AND SPECIAL PURPOSE COMPUTING	DATA MANAGEMENT	THEORETICAL COMPUTER SCIENCE	NETWORK AND RESEARCH FACILITIES	PERF EVAL AND MODEL	TOTAL FUNDING
DOD										
DARPA	-0-	2.8	13.5	11.5	15.8	1.5	1.8	14.0	-0-	60.9
ARMY										
CECOM	-0-	-0-	1.0	0.5	-0-	1.0	-0-	-0-	-0-	2.5
ARO	0.4	0.5	0.7	0.3	0.4	-0-	0.3	-0-	-0-	2.6
AIR FORCE										
AFOSR	2.1	0.8	1.7	0.7	0.2	0.2	0.6	-0-	0.2	6.5
RADC	-0-	1.4	1.2	5.1	-0-	0.2	-0-	0.2	-0-	8.1
AFAL	-0-	0.7	-0-	-0-	-0-	0.6	-0-	0.1	-0-	1.4
NAVY										
ONR	1.4	1.0	2.2	1.0	1.0	-0-	-0-	-0-	-0-	6.6
NAVELEX	-0-	1.7	0.1	0.8	1.0	0.6	0.1	0.7	-0-	5.0
TOTAL DOD	3.9	8.9	20.4	19.9	18.4	4.1	2.8	15.0	0.2	93.6
NSF	0.8	4.3	2.6	2.8	2.2	1.6	6.3	12.2	5.0	37.8
NASA	2.4	4.1	3.7	0.5	3.5	3.6	-0-	1.1	1.4	20.3
DOC	1.2	-0-	1.5	-0-	-0-	-0-	-0-	5.0	0.1	7.8
DOE	7.5	2.7	0.6	1.5	-0-	0.7	0.1	0.8	-0-	13.9
TOTAL	15.8	20.0	28.8	24.7	24.2	10.0	9.2	34.1	6.6	173.4

TABLE B-2
ADVANCED COMPUTING RESEARCH AREA FUNDING SUMMARY BY ORGANIZATION

FY 1984 ACTUAL
(CURRENT DOLLARS, IN MILLIONS)

	COMPUTATIONAL MATHEMATICS	MACHINE ARCHITECTURE (H/W & S/W)	MACHINE INTELLIGENCE AND ROBOTICS	DISTRIBUTED COMPUTING AND SOFTWARE SYSTEMS	VLSI DESIGN AND SPECIAL PURPOSE COMPUTING	DATA MANAGEMENT	THEORETICAL COMPUTER SCIENCE	NETWORK AND RESEARCH FACILITIES	PERF EVAL AND MODEL	TOTAL FUNDING
DOD										
DARPA	-0-	12.2	21.6	14.7	17.8	1.5	2.0	20.3	-0-	90.1
ARMY										
CECOM	-0-	-0-	1.0	2.0	-0-	1.0	-0-	2.0	-0-	6.0
ARO	0.5	0.5	3.3	0.9	0.5	-0-	0.3	-0-	-0-	6.0
AIR FORCE										
AFOSR	2.5	0.8	1.7	0.7	0.3	0.2	0.7	-0-	0.4	7.3
RADC	-0-	1.2	2.0	8.2	-0-	0.6	-0-	1.1	-0-	13.1
AFAL	-0-	0.9	-0-	-0-	0.3	0.4	-0-	0.1	-0-	1.7
NAVY										
ONR	1.5	1.0	2.3	1.3	1.1	-0-	-0-	-0-	-0-	7.2
NAVELEX	-0-	1.5	0.6	0.9	1.2	-0-	-0-	1.9	-0-	6.1
NAVMAT	-0-	-0-	-0-	3.0	-0-	-0-	-0-	-0-	-0-	3.0
TOTAL DOD	4.5	18.1	32.5	31.7	21.2	3.7	3.0	25.4	0.4	140.5
NSF	0.9	3.9	3.9	3.7	2.3	2.3	8.8	14.4	3.8	44.0
NASA	2.2	3.9	3.6	2.0	3.5	1.9	-0-	1.2	1.2	19.5
DOC	1.2	-0-	1.5	-0-	-0-	-0-	-0-	5.0	0.5	8.2
DOE	8.0	2.9	0.5	1.5	-0-	0.7	0.1	0.9	0.1	14.7
TOTAL	16.8	28.8	42.0	38.9	27.0	8.6	11.9	46.9	6.0	226.9

TABLE B-3

ADVANCED COMPUTING RESEARCH AREA FUNDING SUMMARY BY ORGANIZATION

FY 1985 ESTIMATED
(CURRENT DOLLARS, IN MILLIONS)

	COMPUTATIONAL MATHEMATICS	MACHINE ARCHITECTURE (H/W & S/W)	MACHINE INTELLIGENCE AND ROBOTICS	DISTRIBUTED COMPUTING AND SOFTWARE SYSTEMS	VLSI DESIGN AND SPECIAL PURPOSE COMPUTING	DATA MANAGEMENT	THEORETICAL COMPUTER SCIENCE	NETWORK AND RESEARCH FACILITIES	PERF EVAL AND MODEL	TOTAL FUNDING
DOD										
DARPA	-0-	20.4	34.1	17.8	19.6	2.4	3.1	26.7	-0-	124.1
ARMY										
CECOM	-0-	-0-	1.5	3.0	-0-	1.0	-0-	0.5	-0-	6.0
ARO	0.5	0.6	3.6	1.0	0.6	-0-	0.6	-0-	-0-	6.9
AIR FORCE										
AFOSR	3.2	0.6	2.1	0.7	0.3	0.4	0.7	-0-	0.2	8.2
RADC	-0-	1.3	2.6	7.9	0.1	0.4	-0-	0.3	-0-	12.6
AFAL	-0-	0.8	-0-	-0-	1.0	-0-	-0-	-0-	-0-	1.8
NAVY										
ONR	1.6	1.2	2.4	1.3	1.1	-0-	-0-	-0-	-0-	7.6
NAVELEX	-0-	1.6	0.6	0.9	1.2	0.3	-0-	2.0	0.1	6.7
NAVMAT	-0-	0.4	-0-	8.8	-0-	-0-	-0-	-0-	-0-	9.2
SUPERCOMPUTING RESEARCH CENTER	—	12.0	—	—	—	—	—	—	—	12.0
TOTAL DOD	5.3	38.9	46.9	41.4	23.9	4.5	4.4	29.5	0.3	195.1
NSF	1.0	3.7	5.1	4.5	2.4	3.1	11.3	16.0	4.0	51.1
NASA	1.4	4.6	7.0	0.9	2.0	1.7	-0-	2.2	1.1	20.9
DOC	1.2	0.0	1.6	-0-	-0-	-0-	-0-	5.0	1.0	8.8
DOE	9.0	8.0	1.0	2.2	-0-	0.9	0.3	1.0	-0-	22.4
TOTAL	17.9	55.2	61.6	49.0	28.3	10.2	16.0	53.7	6.4	298.3

APPENDIX C

VERY HIGH-PERFORMANCE COMPUTING RESEARCH AREA FUNDING SUMMARY BY ORGANIZATION FISCAL YEARS 1983, 1984, and 1985

TABLE C-1

VERY HIGH-PERFORMANCE COMPUTING
RESEARCH AREA FUNDING SUMMARY BY ORGANIZATION

FY 1983 - ACTUAL
(CURRENT DOLLARS, IN MILLIONS)

	COMPUTATIONAL MATHEMATICS	MACHINE ARCHITECTURE (H/W & S/W)	MACHINE INTELLIGENCE AND ROBOTICS	DISTRIBUTED COMPUTING AND SOFTWARE SYSTEMS	VLSI DESIGN AND SPECIAL PURPOSE COMPUTING	DATA MANAGEMENT	THEORETICAL COMPUTER SCIENCE	NETWORK AND RESEARCH FACILITIES	PERF EVAL AND MODEL	TOTAL FUNDING
DOD										
DARPA	-0-	2.8	0.3	1.1	1.2	-0-	-0-	0.5	-0-	5.9
ARMY										
CECOM	-0-	-0-	-0-	-0-	-0-	-0-	-0-	-0-		-0-
ARO	0.4	0.2	0.2	0.1	0.1	-0-	0.2	-0-	-0-	1.2
AIR FORCE										
AFOSR	1.0	0.6	0.4	0.2	0.1	0.1	0.1	-0-	-0-	2.5
RADC	-0-	-0-	-0-	-0-	-0-	-0-	-0-	-0-	-0-	-0-
AFAL	-0-	0.7	-0-	-0-	-0-	0.6	-0-	0.1	-0-	1.4
NAVY										
ONR	1.4	1.0	2.2	1.0	1.0	-0-	-0-	-0-	-0	6.6
NAVELEX	-0-	1.7	-0-	0.2	-0-	0.2	-0-	-0-	-0-	2.1
TOTAL DOD	2.8	7.0	3.1	2.6	2.4	0.9	0.3	0.6	-0-	19.7
NSF	-0-	3.4	-0-	2.0	0.7	0.1	0.6	0.7	0.1	7.6
NASA	0.4	3.7	-0-	-0-	-0-	-0-	-0-	-0-	0.2	4.3
DOC	-0-	-0-	-0-	-0-	-0-	-0-	-0-	-0-	0.1	0.1
DOE	1.5	2.7	0.3	0.3	-0-	0.1	-0-	-0-	-0-	4.9
TOTAL	4.7	16.8	3.4	4.9	3.1	1.1	0.9	1.3	0.4	36.6

TABLE C-2

VERY HIGH-PERFORMANCE COMPUTING
RESEARCH AREA FUNDING SUMMARY BY ORGANIZATION

FY 1984 ACTUAL
(CURRENT DOLLARS, IN MILLIONS)

	COMPUTATIONAL MATHEMATICS	MACHINE ARCHITECTURE (H/W & S/W)	MACHINE INTELLIGENCE AND ROBOTICS	DISTRIBUTED COMPUTING AND SOFTWARE SYSTEMS	VLSI DESIGN AND SPECIAL PURPOSE COMPUTING	DATA MANAGEMENT	THEORETICAL COMPUTER SCIENCE	NETWORK AND RESEARCH FACILITIES	PERF EVAL AND MODEL	TOTAL FUNDING
DOD										
DARPA	-0-	12.2	1.3	1.2	4.8	-0-	-0-	1.9	-0-	21.4
ARMY										
CECOM	-0-	-0-	-0-	-0-	-0-	-0-	-0-	-0-	-0-	-0-
ARO	0.4	0.3	0.5	0.2	0.2	-0-	0.2	-0-	-0-	1.8
AIR FORCE										
AFOSR	1.3	0.6	0.4	0.2	0.1	0.1	0.1	-0-	-0-	2.8
RADC	-0-	-0-	-0-	-0-	-0-	-0-	-0-	-0-	-0-	-0-
AFAL	-0-	0.9	-0-	-0-	0.3	0.4	-0-	0.1	-0-	1.7
NAVY										
ONR	1.5	1.0	2.3	1.3	1.1	-0-	-0-	-0-	-0-	7.2
NAVELEX	-0-	0.1	-0-	0.4	0.2	0.1	-0-	-0-	-0-	0.8
NAVMAT	-0-	-0-	-0-	-0-	-0-	-0-	-0-	-0-	-0-	-0-
TOTAL DOD	3.2	15.1	4.5	3.3	6.7	0.6	0.3	2.0	-0-	35.7
NSF	-0-	2.9	-0-	2.5	1.4	0.1	0.8	1.7	0.1	9.5
NASA	0.6	3.5	-0-	-0-	-0-	-0-	-0-	-0-	0.2	4.3
DOC	-0-	-0-	-0-	-0-	-0-	-0-	-0-	-0-	0.5	0.5
DOE	4.0	2.9	0.3	0.4	-0-	0.2	-0-	-0-	-0-	7.8
TOTAL	7.8	24.4	4.8	6.2	8.1	0.9	1.1	3.7	0.8	57.8

TABLE C-3

VERY HIGH-PERFORMANCE COMPUTING
RESEARCH AREA FUNDING SUMMARY BY ORGANIZATION

FY 1985 ESTIMATED
(CURRENT DOLLARS, IN MILLIONS)

	COMPUTATIONAL MATHEMATICS	MACHINE ARCHITECTURE (H/W & S/W)	MACHINE INTELLIGENCE AND ROBOTICS	DISTRIBUTED COMPUTING AND SOFTWARE SYSTEMS	VLSI DESIGN AND SPECIAL PURPOSE COMPUTING	DATA MANAGEMENT	THEORETICAL COMPUTER SCIENCE	NETWORK AND RESEARCH FACILITIES	PERF EVAL AND MODEL	TOTAL FUNDING
DOD										
DARPA	-0-	23.8	2.5	1.3	5.3	1.0	0.6	3.5	-0-	38.0
ARMY										
CECOM	-0-	-0-	-0-	-0-	-0-	-0-	-0-	-0-	-0-	-0-
ARO	0.4	0.4	0.6	0.3	0.2	-0-	0.3	-0-	-0-	2.2
AIR FORCE										
AFOSR	1.6	0.5	0.5	0.2	0.2	0.1	0.2	-0-	-0-	3.3
RADC	-0-	-0-	-0-	-0-	-0-	-0-	-0-	-0-	-0-	-0-
AFAL	-0-	0.8	-0-	-0-	1.0	-0-	-0-	-0-	-0-	1.8
NAVY										
ONR	1.6	1.2	2.4	1.3	1.1	-0-	-0-	-0-	-0-	7.6
NAVELEX	-0-	0.4	-0-	0.6	0.6	0.2	-0-	-0-	-0-	1.8
NAVMAT	-0-	-0-	-0-	-0-	-0-	-0-	-0-	-0-	-0-	-0-
SUPERCOMPUTING RESEARCH CENTER	-0-	12.0	-0-	-0-	-0-	-0-	-0-	-0-	-0-	12.0
TOTAL DOD	3.6	39.1	6.0	3.7	8.4	1.3	1.1	3.5	-0-	66.7
NSF	-0-	3.3	0.2	3.0	1.7	0.2	1.0	2.0	0.2	11.6
NASA	0.8	4.4	-0-	-0-	-0-	0.6	-0-	-0-	0.1	5.9
DOC	-0-	-0-	-0-	-0-	-0-	-0-	-0-	-0-	1.0	1.0
DOE	6.3	8.0	0.5	0.6	-0-	0.3	-0-	-0-	-0-	15.7
TOTAL	10.7	54.8	6.7	7.3	10.1	2.4	2.1	5.5	1.3	100.9

APPENDIX D

FEDERAL RESEARCH ACTIVITIES IN COARSE- AND FINE-GRAIN ARCHITECTURES

TABLE D-1

FEDERAL RESEARCH ACTIVITIES IN
COARSE-GRAIN ARCHITECTURES
(CHARACTERIZED BY INTERCONNECTION NETWORK)

PROGRAM	SPONSOR	PERFORMER
PACKET SWITCHING		
Ultracomputer	DOE/NSF	NYU
CEDAR	DOE/NSF	Illinois
FMP	NASA	Stanford
Homogeneous Machines	DARPA/NASA/DOE	Cal Tech
PUMPS	NSF	Purdue
Static Dataflow	NASA/DOE/NSF	MIT
Tagged Token Dataflow	DARPA	MIT
Ring Dataflow	NASA/DOE	Lawrence Livermore
CIRCUIT SWITCHING		
TRAC	NSF/DOE/AFOSR	Texas
Butterfly	DARPA/NSF	BBN
TREE		
DADO	DARPA	Columbia
AMPS	NSF	Utah
Reduction Tree	NSF	UNC
NEAREST NEIGHBOR		
Finite Element Machine	NASA	NASA/Langley
Navier Stokes Machine	NASA	Princeton
Wavefront Array	NASA/ONR	USC
DAISY IV	NSF	USC
Cosmic Cube	DOE/DARPA	Cal Tech
CROSSBAR		
Database Machine	DOE/NSF	Wisconsin
Multi-Micros	DOE	Los Alamos
S-1	NAVELEX	Lawrence Livermore
RING		
ZMOB	NSF/AFOSR	Univ. of Maryland
Psuedo-Ring	NSF	UC, Santa Barbara
Crystal	NSF	Wisconsin
MISC		
PASM	ONR	Purdue
MD/C	ONR	Princeton
Parallel Speech	NSF	Purdue
Special Purpose Array Processor	NSF	Northwestern
Special Purpose Array Processor	NSF	Ohio State Univ.
Speech Architecture	NSF	Brown

TABLE D-2

FEDERAL RESEARCH ACTIVITIES IN
FINE-GRAIN ARCHITECTURES

PROGRAM	SPONSOR	PERFORMER
Connection Machine	DARPA	Thinking Machines
Boolean Vector Machine	NSF	Duke University
Massivelly Parallel Processor	NASA/Goddard	Goodyear
Non-Von	DARPA	Columbia
Blue Chip	ONR	Purdue University
Programmable Systolic Array	DARPA	Carnegie-Mellon University
Pixel Planes Processor	DARPA/NSF	University of North Carolina
Pipeline Bit-Serial Processor Array	NSF	Duke

APPENDIX E

ADDITIONAL CONTRIBUTORS TO THE REPORT

ADDITIONAL CONTRIBUTORS TO THE REPORT

A. UNITED STATES ARMY

Army Research Office
Dr. Jimmie Suttle
CENTACS
Dr. James Schell

B. UNITED STATES NAVY

Naval Electronics Systems Command
Mr. John Machado
Office of Naval Research
Dr. Paul Schneck

C. UNITED STATES AIR FORCE

AFOSR
Dr. David Fox
Major John Thomas
Avionics Laboratory
Dr. Donald Moon
Rome Air Development Center
Colonel David Carlstrom
Mr. Sam DiWitte
Mr. Richard Metzger
Systems Command
Lietuenant Colonel James Riley

D. DEFENSE ADVANCED RESEARCH PROJECTS AGENCY

Dr. Craig I. Fields

E. DEPARTMENT OF ENERGY

Dr. James Decker

F. NATIONAL SCIENCE FOUNDATION

Mr. Kent Curtis

ARGONNE NATIONAL LABORATORY
9700 South Cass Avenue
Argonne, Illinois 60439

ADVANCED ARCHITECTURE COMPUTERS

Jack J. Dongarra and Iain S. Duff

Mathematics and Computer Science Division

Technical Memorandum No. 57

October 11, 1985

TABLE OF CONTENTS

Advanced Architecture Computers*

Jack J. Dongarra and Iain S. Duff

(dongarra@anl-mcs and na.duff@su-score)

Mathematics and Computer Science Division
Argonne National Laboratory
Argonne, Illinois 60439

Abstract We describe the characteristics of several recent computers that employ vectorization or parallelism to achieve high performance in floating-point calculations. We consider both top-of-the-range supercomputers and computers based on readily available and inexpensive basic units. In each case we discuss the architectural base, novel features, performance, and cost. It is intended that this report will be continually updated, and to this end the authors welcome comment.

Keywords

vector processors, array processors, parallel architectures, supercomputers, high-performance computers

1. Introduction

In the last few years several machines have been announced that use some form of parallelism to achieve a performance in excess of that attainable directly from the underlying technology used in the design of the constituent chips. To a large degree the availability of low-cost chips as building blocks has given rise to many of these new machines. We give a list of the chips so used in Appendix A.

After listening to a great number of both technical and sales presentations on these new computers, we quickly became overwhelmed and confused with the characteristics of each product and its relative strengths and weaknesses. In an effort to clarify our understanding, we have written this report summarizing the principal features of each machine. We hope that the publication of this report will provide similar assistance to other computational scientists and will clarify what

*Work supported in part by the Applied Mathematical Sciences subprogram of the Office of Energy Research, U. S. Department of Energy, under Contract W-31-109-Eng-38. During preparation of this report, the second author was on leave from CSS Division, AERE Harwell, Didcot, OXON OX11 0RA, England. This draft was typeset on October 11, 1985.

architectures are currently being employed and the range of machines available.

In Section 2 we list the computers considered and discuss the criteria we have used to select these computers. We present a rough classification based on architectural features and use this in our list of machines. We also summarize principal features of the machines in two tables: one for the expensive supercomputers and the other for cheaper machines. More detailed information on the machines is provided as Appendix B of this report.

The guidelines used in preparing the detailed descriptions are given in Section 3. In some cases, our data are incomplete and nonuniform. This situation reflects the technical level of the presentations, the documentation available to us, and the stage of development of the product being described. We would be grateful for comments and criticisms that might help to remedy these deficiencies. We intend to update this report from time to time to reflect both the changing marketplace and further information on currently listed machines.

2. Summary and Classification of Machines Considered

In recent months there has been an unprecedented explosion in the number of computers in the marketplace. This explosion has been fueled partly by the availability of powerful and cheap building blocks and by the availability of venture capital. There have been two main directions to this explosion. One has been the personal computer market and the other the development and marketing of computers using advanced architectural concepts. In this report we restrict our study to the latter group, with particular interest in architectures that use some form of parallelism to increase performance over that of the basic chip.

We also restrict our attention to machines that are available commercially, and thus exclude research projects in universities and government laboratories and products still at the design stage. For the sake of completeness (and to solicit further comment from the academic community) we do, however, list in Appendix C those experimental machines with which we are familiar. We have included projects only if they are still active; several experimental projects are moribund or have been superseded by commercial exploitation. We have excluded some machines that we have heard about only second- or thirdhand. Thus there are many experimental machines that we have not listed. We would, however, be delighted to be alerted to ongoing activities.

Some machines not commonly thought of as multiprocessors can be used as such. For example, the IBM 3081 and 3084 are multiple-processor machines. Most installations use this feature to increase the throughput, but it is possible to use

them as multiple processors (with multiplicity up to 2 and 4 for the two machiines, respectively) using the IBM Program Product MTF which runs under MVS. We do not, however, give further details of these machines in Appendix B. We also do not include details on attached processors (for example, the Univac AP) since they are much more limited in their applicability and could be considered more as special-purpose engines which assuredly fall outside the scope of this paper.

We have also necessarily had to exclude information obtained under non-disclosure agreements. We will update this report as such information is released through product announcements.

A much-referenced and useful taxonomy of computer architectures was given by Flynn(1966). He divided machines into four categories:

(i) SISD - single instruction stream, single data stream
(ii) SIMD - single instruction stream, multiple data stream
(iii) MISD - multiple instruction stream, single data stream
(iv) MIMD - multiple instruction stream, multiple data stream

Although these categories give a helpful coarse division, we find immediately on examining current machines that the situation is more complicated, with some architectures exhibiting aspects of more than one category.

Many of today's machines are really a hybrid design. For example, the CRAY X-MP has up to four processors (MIMD), but each processor uses pipelining (SIMD) for vectorization. Moreover, where there are multiple processors, the memory can be local, global, or a combination of these. There may or may not be caches and virtual memory systems, and the interconnections can be by crossbar switches, multiple bus-connected systems, time-shared bus systems, etc.

With this caveat on the difficulty of classifying machines, we list below the machines considered in this report. We group those with similar architectural features. We have not included the four machines Axiom, Culler, Vitesse, and Mips in this list since the documentation we have on these machines has insufficient technical details for us to classify them.

scalar
- pipelined (e.g., 7600, 3090)
- parallel pipelined microcoded
  - *FPS 164*
  - *FPS 264*
  - *Multiflow*
  - *STAR ST-100*

vector
- memory to memory
  - *CDC CYBER 205*
- register to register
  - *American Super.*
  - *Convex C-1*
  - *CRAY-1*
  - *CRAY X-MP-1*
  - *Amdahl 500, 1100, 1200, 1400 (Fujitsu VP-50, 100, 200, 400)*
  - *Galaxy YH-1*
  - *Hitachi S-810*
  - *NEC SX-1, SX-2*
  - *Scientific Computer Systems*
- cache based r-to-r
  - *Alliant FX/1*

parallel
- global memory
  - bus connect
    - *Alliant FX/8 (vector capability)*
    - *Elxsi 6400*
    - *Encore Multimax*
    - *Flex/32*
    - *IP-1*
    - *Sequent Balance 8000*
  - direct connect
    - *Amer. Super. (vector capability)*
    - *CRAY-2 (vector capability)*
    - *CRAY-3 (vector capability)*
    - *CRAY X-MP-2/4 (vector cap.)*
    - *Denelcor HEP-1*
- local memory
  - hypercube
    - *Ametek System 14*
    - *Intel iPSC*
    - *NCUBE*
    - *Connection Machine*
  - butterfly
    - *BBN Butterfly*
  - ring-bus
    - *CDC CYBERPLUS*
  - lattice
    - *Goodyear MPP*
    - *ICL DAP*
  - dataflow
    - *Loral DATAFLO*
- multilevel memory
  - *ETA-10 (vector capability)*
  - *Myrias 4000*

A more empirical subdivision can be made on the basis of cost. We split the machines into two classes: those costing over $1 million and those under $1 million. The former group is usually classed as supercomputers, the latter as high-performance engines. With this subdivision, we can summarize the machines in the following tables. Since we do not have sufficient technical information on the PS 2000, the Galaxy YH-1, and the Culler, Vitesse, and Mips machines, we have excluded them from these summary tables.

Table 1
Machines Costing over $1M (base system)

Machine	Word length	Maximum Rate in MFLOPS	Memory in Mbytes	OS	Number of Proc.
Amdahl 1400 (Fujitsu VP-400)	32/64	1142	256	Own	1
CRAY-1	64	160	32	Own	1
CRAY X-MP	64	210/proc	128	Own/UNIX	1,2,4
CRAY-2	64	500/proc	2000	UNIX	4
CRAY-3	64	1000/proc	16000	UNIX	16
CYBER 205	32/64	400	32	Own	1
CYBERPLUS	32/64	100/proc	4(a)	Own	256
Denelcor HEP-1	32/64	10/PEM	16/PEM	UNIX	16(b)
ETA-10	32/64	1250/proc	2048(c)	Own	2,4,6,8
Hitachi S-810/20	32/64	840	256	Own	1
Myrias 4000	32/64/128	???	512/Krate	UNIX	1024/Krate
NEC SX-2	32/64	1300	256(d)	Own	1

(a) Memory per processor
(b) 64 processes possible for each PEM; however, effective parallelism per PEM is 8-10.
(c) Also 32 Mwords of local memory with each processor
(d) Also a 2-Gbyte extended memory

The actual price of the systems in Table 1 is very dependent on the configuration, with most manufacturers offering systems in the $5 million to $20 million range. All use ECL logic with LSI (except the CRAY-1 in SSI, CRAY X-MP, and HEP in MSI), and all use pipelining and/or multiple functional units to achieve vectorization/parallelization within each processor. For the multiple-processor

systems, the form of synchronization varies: event handling on the CRAYs, asynchronous variables on the HEP, send/receive on the CYBERPLUS. The CRAY-3 and ETA-10 are not yet available. Both Amdahl and Hitachi systems are IBM System 370 compatible.

In Table 2 we summarize machines in the lower price category.

Table 2
Machines costing under $1M

Machine	Chip	Parallelism	Connection
Alliant FX/8	WTL 1064/1065 plus 10 gate arrays	8+vector	cross bar (reg to cache) and bus (cache to memory)
Ametek System 14	80286/80287	256	hypercube
Amer. Super. Comp.	ECL	Vector	(vector)
Axiom	LSI	1	(scalar)
BBN Butterfly	68020/68881	256	butterfly
Think Machines/ Connection	VLSI	64000	hypercube
Convex C-1	Gate array	Vector	(vector)
Elxsi 6400	ECL	12	bus
Encore Multimax	32032/32081	20	bus
Flex/32	32032/32081	20	bus
FPS 364	LSI	1	(scalar)
FPS 264	ECL	1	(scalar)
FPS 164+MAX	VLSI	16	bus
FPS 5000	VLSI	4	bus
FPS MP32	VLSI	3	bus
ICL DAP	ECL	1024	near-neighbor
Intel iPSC	80286/80287	128	hypercube
IP-1	????	8	cross-bar
Loral DATAFLO	32016	128	bus
Goodyear MPP	VLSI	16384	near-neighbor
Multiflow	gate array	8	(scalar)
NCUBE	Custom VLSI	1024	hypercube
SCS-40	ECL/LSI	Vector	(vector)
Sequent Balance 8000	32032/32081	12	bus
Star ST-100	VLSI	1	(scalar)

The data presented in Table 2 differ from that of Table 1. Full details for all the machines are given in Appendix B. Because of the widely differing architectures of the machines in Table 2 it is not really advisable to give one or even two values for the memory. In some instances there is an identifiable global memory; in others there is a fixed amount of memory per processor. Additionally, it may be possible to configure

memory either as local or global. A value for the maximum speed is even less meaningful than in Table 1, since a high Megaflop rate is not necessarily the objective of the machines in Table 2 and the actual speed will be very dependent upon the algorithm and application. In the other aspects quoted in Table 1, all the machines in Table 2 are very similar. All machines, except the FPSs and the SCS (all 64 bit), the DAP, MPP, and Connection (all bit-slice, the first two supporting variable-precision floating point), and the Star (32 bit), have both 32- and 64-bit arithmetic hardware, with most of them adhering closely to the IEEE standard. The Loral and IP-1 machines are not designed for floating-point computation. Also, all machines have a version of UNIX as their operating system, except FPS (host system), American Supercomputer and SCS (COS), and Star and Ametek (own system).

3. Template for Machine Description

As we mentioned in the introduction, the level of technical information on each machine varied significantly. We have, however, attempted to organize the available information in a consistent manner. In Table 3, we give the template used in presenting the data in the appendices.

Reference

Flynn, M. J. (1966) Very high-speed computing systems. Proc IEEE, vol. 54, pp. 1901-1909.

Table 3
Template for Description of Machines

Name of machine, manufacturer, backers, etc.
Contact: technical and sales
Architecture
  Basic chip used
  Local, global-shared memory, or both
  Connectivity (for example, grid, hypercube)
  Range of memory sizes available; virtual memory
  Floating point unit (IEEE standard?)
Configuration
  Stand-alone or range of front-ends
  Peripherals
Software
  UNIX or other?
Languages available
Fortran characteristics
  F77
  Extensions
  Debugging facilities
  Vectorizing/parallelizing capabilities
Applications
  Run on prototype
  Software available
Performance
  Peak
  Benchmarks on codes and kernels
Status
  Date of delivery of first machine, beta sites, etc.
  Expected cost (cost range)
  Proposed market (numbers and class of users)

APPENDIX A

LIST OF BASIC CHIPS USED

General-Purpose Floating-Point Processors

Intel 8087/80287
National 32081
Motorola 68881
Zilog 8070
AMD 9511A/9512
Fairchild F9450

Building-Block Floating-Point Processors

Weitek WTL1032/1033
TRW TDC 1022/1042
Weitek WTL 1064/1065
AMD 29325
Analog Devices ADSP2310/2320

General-Purpose Building-Block Floating-Point Processors

Weitek WTL 1164/1165 (Fandrianto and Woo 1985)

Reference

Fandrianto, J. and Woo, B.Y. (1985), VLSI floating-point processors. IEEE Proceedings of the 7th Symposium on Computer Arithmetic, pp. 93-100.

APPENDIX B

DETAILS OF MACHINES CONSIDERED

ALLIANT FX/1 and ALLIANT FX/8

Alliant Computer Systems Corp.
42 Nagog Park
Acton, MA 01720

617-263-9110

Contact: Technical: Craig J. Mundie, vice president of software
Contact: Sales: David L. Micciche, vice president marketing, sales
and customer servies

Backers: Venrock
Hambrecht and Quist
Kleiner, Perkins, Caulfield and Byers

Formerly, the company was called Dataflow.

Vector Register Parallel Shared Memory Architecture

Computational elements (CEs) execute applications code using vector instructions. An FX/1 has one CE. An FX/8 has 1-8 CEs. The CEs transparently execute the code of an application in parallel. CEs may be added in the field, increasing performance without recompilation or relinking.

Each CE has 8 vector registers, each with 32 64-bit elements, and 8 64-bit scalar floating point, 8 32-bit integer, and 8 32-bit address registers.

Interactive Processors (IPs) execute operating system, interactive code, and I/O operations. An FX/1 has 1-2 IPs. An FX/8 has 1-12 IPs.

Basic chip used: Weitek 1064/1065 plus ten different gate array types with 2600 to 8000 gates. In addition, the Motorola 68012 is used in the IP. The cycle time is 170 ns.

CEs are cross-bar connected on the backplane to a 64 Kbyte/128 Kbyte write-back computational processor (CP) cache (FX/8). Bandwidth is 376 Mbyte/sec.

Each 32-Kbyte IP cache is connected to 1-3 IPs (FX/8) or 1-2 IPs and a CE (FX/1). The FX/8 has 1-4 IP caches; the FX/1 has one IP cache.

The CP and IP caches are attached by two 72-bit busses to the main memory. Memory bus bandwidth is 188 Mbyte/sec.

Connectivity: crossbar (CE to cache), bus (cache to memory, cache to cache)

Range of memory sizes available: 8-16 Mbytes (FX/1), 8-64 Mbytes (FX/8), all with ECC.

Virtual memory: 2 Gbytes per process

Floating point unit: IEEE 32- and 64-bit formats including hardware divide and square root and microcoded elementary functions.

Configuration: Standalone. TCP/IP network support.

Size (inches): FX/1 system - 28h x 13w x 25d (the FX/1 I/O expansion cabinet is the same size); FX/8 system - 43h x 29w x 34d (the FX/8 I/O expansion cabinet is 22w and same height and depth).

Cooling: Both the FX/8 and FX/1 are air-cooled. The FX/8 system consumes 4950 watts (max. configuration), the FX/1 system 1155 watts (max. configuration).

Peripherals:

800/1600/6250 BPI start-stop tape drive
67, 134, and 379 Mbyte (formatted) Winchester disk drives
45 MBbyte cartridge tape drive
Floppy disk drive
8/16 line multichannel communications controllers
600 lpm printer
Ethernet controller

Software: Concentrix, Alliant's enhancement of Berkeley 4.2 UNIX with multiprocessor support. Compiler runs on production hardware and software.

Languages: Fortran, C, Pascal

Fortran characteristics:

F77 - Conforms to 1978 ANSI standard.
Extensions - Most of VAX/VMS extensions and Fortran 8x array extensions.
Debugging facilities - Yes.

Vectorizing/parallelizing capabilities - Automatic detection of vectors and parallelism.
Feedback to user via diagnostic messages.
User control of transformations via directives in the form of Fortran comments.

Performance:

Scalar 32 bit - 4.45 MIPs / CE. (4450 Kwhetstones)
Scalar 64 bit - 3.63 MIPs / CE. (3630 Kwhetstones)
Vector 32 bit: 11.8 MFLOPS / CE.
(1 chime multiply-add triad at 170ns/chime)
Vector 64 bit: 5.9 MFLOPS / CE.
(2 chime multiply-add triad at 170ns/chime)
(64-bit multiply is 2 chimes; 64-bit add, subtract, and move are 1 chime).

Applications: Engineering and scientific end-user and OEM applications, stand-alone or as a computational server to a network of engineering workstations.

Status: First beta delivery May 1985; first production shipment September 1985.

Expected cost: FX/1 - $132,000 to $200,000; FX/8 - $270,000 to $750,000

Amdahl Vector Processors (Fujitsu VP)

John Roberts
Amdahl Corp.
1250 East Arques Ave.
P.O. Box 3470
Sunnyvale, CA 94088

408-746-6880

Vector Register Architecture

The Amdahl 500, 1100, 1200 and 1400 Vector Processors are marketed by Amdahl Corp. in the U.S., Canada, and Europe. These products are manufactured by Fujitsu, and similar models are marketed in Japan as the VP-50, VP-100, VP-200, and VP-400. The VP-100 and 200 is also marketed by Siemens in mainland Europe.

These are all register-to-register machines. All models have one scalar and one vector unit which can execute computations independently. The scalar unit fetches all instructions and passes each instruction to the appropriate unit for execution. The scalar processor is based on the Fujitsu M380/382 series mainframes and runs 185 of the IBM S/370 instruction set plus 10 unique instructions. The vector performance varies according to model as follows:

Model	Peak MFLOPS
500	133
1100	267
1200	533
1400	1142

The scalar processor cycle time is 14 ns (VP 1400 only) or 15 ns (compared to the X-MP's 9.5 ns), but a sampling of scalar instructions indicates that the VP operations may be slightly faster than the X-MP's. There is, moreover, a difference in the pipelining between the X-MP and VP. Each VP scalar instruction is pipelined in three stages: fetch, decode, and execute. However, unlike the X-MP, the execution stage in the VP is not segmented. Thus, there is less potential purely scalar overlap in the VP than in the X-MP. (Note that all scalar work can overlap vector operations.)

The vector unit consists of 5 or 6 pipelines, a vector register memory, and a mask memory. The 5 or 6 pipelines comprise 1 or 2 load/store pipelines, plus 1 mask pipeline, 1 add/logical pipeline, 1 multiply pipeline, and 1 divide pipeline. The number of concurrent pipelines, vector register size, and mask register size differ for each model, as shown below. Main memory capacity ranges from 32 Mbytes to 256 Mbytes (4 to 32 Mwords).

Configuration	500	1100	1200	1400
# pipes total	5	6	6	5
# concurrent load/store pipes	1	2	2	1
# 64 bit words/vect cyc/pipe	1	1	2	4
Scalar cycle time (ns)	15	15	15	14
Vector cycle time (ns)	7.5	7.5	7.5	7
# concurrent arith pipes	1	2	2	2
# 64-bit results/vect cyc/pipe	1	1	2	4
Vect. reg. size (Kbytes)	32	32	64	128
Mask reg. size (Bytes)	512	512	1024	2048
Max. main memory (Mbytes)	128	128	256	256
Min. main memory (Mbytes)	32	32	64	64
Max interleaving (ways)	128	128	256	256

The total vector register capacity is 32-128 Kbytes. The registers can be reconfigured dynamically to 6 different combinations with varying vector register lengths, as shown below:

Configuration of Vector Registers

Register Length by Model
(# of 64-bit word elements)

# registers	500	1100	1200	1400
8	512	512	1024	2048
16	256	256	512	1024
32	128	128	256	256
64	64	64	128	128
128	32	32	64	64
256	16	16	32	32

Technology:

400 and 1300 gate ECL, 350-picosecond delay
main memory - 64 Kbit, 55 ns, MOS static RAM
380-470 square feet
36-62 KVA power consumption
air cooled

Software:

Automatic vectorizing Fortran compiler
Scalar Fortran compiler
Interactive debugger
Performance measurement tools
Interactive vectorizer
Scientific subroutine library (223 routines)

AMETEK System 14

AMETEK/Computer Research Division
610 North Santa Anita Avenue
Arcadia, CA 91006

Technical Contact: Tim Brown
Sales: Floyd Sherman
818-445-6811

Hypercube Architecture

This is the first generation of AMETEK Concurrent Processing Systems. Each node is based on a 80286/80287 Applications Processor/Floating Point Co-processor with a separate 80186 Communication Processor. Each node has 8 bidirectional communications channels at 3 Mbits/sec connected to the host machine through a 1 Mbyte/sec parallel interface.

Local memory - 1 Mbyte per node.

Connectivity - 16 to 256 nodes are connected in hypercube to form a System 14.

Floating Point Unit - IEEE Standard Floating Point Arithmetic

Configuration: Front-end machines (hosts) are DEC VAXs (Microvax II through VAX 780). Support is available for the host running either UNIX 4.2bsd or VMS. The nodes themselves run the AMETEK Hyperent Operating System, specifically designed to support concurrent processing. HOS supports automatic message buffering, message forwarding, process creation, and machine partitioning for multiple users.

Languages: C and Fortran

Software: Simulator, Multi-Process Dubugger, and User Interface to configure the nodes in ring, nearest neighbor, or hypercube topologies; libraries of matrix routines, Fourier transforms.

Status: Beta site machines being shipped fall 1985, production shipments in January 1986

American Supercomputer Company

American Supercomputer Company
130 Keifer Ct., Suite 200
Sunnyvale, CA 94086
408-720-1010

CEO: Bob O. Evans
COO/President: Thomas F. Hunter
VP Engineering: Michael J. Flynn

Vector Register Architecture

Two modes of operation: binary CRAY compatible and ASI mode (which incorporates ASI extensions to the CRAY X-MP architecture). COS has been encapsulated under UNIX System V.

This is a 64-bit scientific computer, designed as stand-alone processor or with interfaces to NSC hyperchannel, IBM, CRAY, DEC, and Ethernet/TCPIP.

A family of processors, with the entry level system one half of a single CPU X-MP.

Status: Benchmarking will commence the fourth quarter of 1986; beta shipment is scheduled for the first quarter of 1987.

Cost: entry-level system under $1 million.

AXIOM Systems

1589 Centre Pointe
Milpitas, California 95035

Richard Lipes
Bob Rau
408-943-9460

Ross Towle (compiler person, student of Kuck)
Bob Rau (Architech from University of Illinois and Elxsi)

Dataflow Architecture

BBN Butterfly

Bolt, Beranek and Newman

Randy Rettberg
BBN Labs
Cambridge, MA

617-497-3538

Parallel Butterfly Network Architecture

First major computer to emerge from the DARPA Multiprocessor Architecture Program.

Based on the MC 68000 boards with 1 Megabyte of memory connected through a very high performance switch. The switch is capable of supporting up to 256 processors.

Other features:

Tightly coupled, shared memory.
All processors given equal access to memory.
Switched access to memory.
Butterfly switch
  packet-switching techniques
  serial data transfer
  simple switching nodes
  O(n log n) switching nodes
  32-Mbps data rate
  4-microsecond access time

A single-processor node is a MC68000 for processing element, an AMD-2901 bit slice processor for memory management, and a custom-designed VLSI switch circuit. It also includes an 8-MHz clock, from 256 Kbytes to 4 Mbytes of memory, a memory management system, interfaces to the switch, a microcoded controller, and an on-board switching regulator. The memory can be either locally or globally accessed through "Butterfly" (or shuffle) network.

Software: Chrysalis O/S (similar to UNIX). The machine can be segmented to allow other users access.

Languages: C and LISP

Cost: $800,000 for 128-processor system

Connection Machine

Thinking Machines Inc.
245 First St.
Cambridge, Mass. 02142-1214

617 876-1111

James Bailey - Director of Marketing

Parallel Hypercube Architecture

The Connection Machine is a very fine grain parallel computer with an architecture suitable for artificial intelligence applications. The 64000 node processor prototype will have 1000 times the logical inference performance capabilities of current LISP workstations.

The processing elements are one-bit machines having 4096 bits of memory connected so that each processor can communicate with any other through a fast message-routing system that forms a hypercube network. All linkages are software controlled with system-wide message flow being handled by a 3 Gigabit per second message routing system. All memory is dual ported and is hence directly accessible by both the Connection Machine and the front end.

Configuration: The Connection Machine system has 65536 physical processors but may be configured for a much larger number of logical processors by means of the global-reset and configure commands.

Access is through a front-end processor, currently either a VAX or a Symbolics 3600. The front-end provides the operating system environment, including terminal interaction and file management.

The clock rate may range up to 10 MHz, giving an expected performance of 2 billion 32-bit integer additions per second in the 64K (65536) node configuration. Average instruction mixes are expected to exceed 1000 Mips.

I/O can be through the front end or direct to a 1.2 Gigabyte disk at the rate of 500 Megabits per second.

Languages: Applications programs reside in the host and can be written in CM-C (a Connection Machine extension of C), CM-Lisp, or an assembly language REL-2.

Applications: One of the principal applications is expected to be image processing. Other applications include VLSI simulation and FFT's.

Status: The first 64000 processor prototype will be delivered to DARPA before the end of 1985. The prototype available in 1986 will use a conservative VLSI technology of 10000 gate CMOS gate arrays.

Convex C-1

Convex Computer Corporation
1819 Firman, Suite 151
Richardson, Texas 75081

Phone: 214-952-0200

Technical: Steve Wallach
Sales: Bob Shaw

Vector Register Architecture

The machine is based on CMOS VLSI gate array 8000 gates/chip (24 different chips in the machine). It uses vector architecture, register to register, with pipelined functional units (each of which operates asynchronously). The machine is based on a 100-ns major cycle time, 50-ns minor cycle time, with virtual memory (page size 4096 bytes) and 1024 bytes cache between memory and registers (64-Kbyte, 50-ns access).

Physical memory - up to 128 MB dynamic RAM(4-way interleaved).
Virtual address space - 4 Gbytes
User address space - 2 Gbytes.
Memory - on a 16-Mbyte board, 2 banks per board, each 4-way interleaved.
Transfer rates between memory and CPU - rated at 80 Mbytes/sec.
Single memory pipe between memory and registers.
Note: 64-bit vector references that are aligned on 64-bit boundaries will bypass the cache.
Vector registers - 8, each with 128 elements (64-bit elements).

VL and VS registers
0.25 Mbyte IOP buffer
I/O transfer rates of 80 Mbyte/sec

Floating point IEEE Standard format.
5 independent I/O processors each rated at 80 Mbyte/sec.
Concurrent operation of scalar and vector units.
Mask/merge and compress operations supported.
Reduction operators max,min,sum,prod,any,all, and parity supported.
Degradation for indirect addressing not specified.

```
A(i) = B(C(i)) ...
      LD VL
      LD C,V0
      SHF 4,V0,V1
      LD B,V0,V1
      STORE A
```

Byte-addressable with integer*2 and 16-bit arithmetic supported.

Configuration: Designed as a stand-alone multiuser machine.

Software: UNIX 4.2 bsd operating system.

Languages: Fortran 77 and C (accepts VMS Fortran), with excellent Fortran vectorizing compiler. Fortran compiler accepts VAX VMS Fortran.

Performance: Peak performance 20 MFLOPS in double precision (64-bit arithmetic), 40 MFLOPS in single precision (32-bit arithmetic). LINPACK timings - expect around 2-4 MFLOPS.

Note: Convex rates their machine as 1/6 of a CRAY 1-S, 600 ns per subroutine call, 9 cycles latency (cf. 11 on CRAY, 30 on FACOM VP)

Basic system: two 19-in. racks and 4-Mbytes memory, 1 I/O processor, service processor, 414 Mbyte Winchester, 6250 bpi tape drive.

Size: 25 x 62 x 40 inches for each cabinet. Base system requires two cabinets, each about 500 lb.

Forced air cooling.
Power consumption 3200-4500 watts

Cost: base system $500,000

4 Mbytes	414-Mbyte disk	one IOP [16 lines]	$495,000
8 Mytes	828-Mbyte disk	one IOP "	$545,000
16 Mbytes	828-Mbyte disk	one IOP "	$595,000

32 Mbytes	828-Mbyte disk	two IOP [32 lines]	$745,000
64 Mbytes	1656-Mbyte disk	two IOP "	$995,000
128 Mbytes	3312-Mbyte disk	two IOP "	$1,400,000

3312 Mbytes = 8 Fuji eagles
can have 3 asynchronous 16 line ports

F77 compiler	$24.5K
(has GPROFF run-time profiler)	
Networking package	$15K

Status: Three machines built and running at beta sites - United Technology, Boston Division, and Mostec (and in-house). Hope to sell 30 machines in first quarter of sales and to sell to Europe by 1986.

CRAY-1

Cray Research Inc.
1440 Northland Drive
Mendota Heights, MN 55120
612-452 6650

Vector Register Architecture

This machine is no longer being produced, although when first introduced in 1976 (Los Alamos), it was undisputedly the fastest processor in the world and is still used as a benchmark for high-speed computing. Since many CRAY customers are currently upgrading their systems to an X-MP, there are opportunities to buy second hand CRAY-1s at knockdown prices.

Features:

A uni-processor.
Vector processor, uses pipelining and chaining to gain speed.
12.5-nsec clock. Fast scalar.
Uses only four chip types with 2 gates per chip.
64-bit word size up to 4 M words of storage.

The CRAY 1-S has bipolar (in units of 4K RAM), and the newer (1982) CRAY 1-M has MOS memory (in units of 16K RAM).

Logic chips - ECL with a gate delay of .7 nsec.
Main memory banked up to 16 ways. The bank busy time is 50 nsec (70 nsec on 1-M) and the memory access time (latency) is 12 clocks (150 nsec).
No virtual memory
Register-to-register machine
8 registers of length 64 (64-bit) words each
Word addressable (64-bits).
No half precision.
Double precision is through software and is extremely slow (factors of about 50 times single precision are common).

There is only one pipe from memory to vector registers, resulting in a major bottleneck with loads and stores to memory from registers. Loads can be chained

with arithmetic operations; stores cannot.

Performance:

Low vector startup times and fast scalar performance make this a very general-purpose machine. Max. performance 160 MFLOPS; 64-bit arithmetic; max. attainable sustained performance 150 MFLOPS. There are codes for matrix multiplication and the solution of equations which get close to this. Maximum scalar rate is 80 MIPS. It is easy to attain over 100 MFLOPS for certain problems, even using Fortran.

Software:

An extensive range of software exists for this machine. Since the instruction set is compatible with the X-MP range, this software will also run on that range.

CRAY-2

Cray Research Inc.
1440 Northland Drive
Mendota Heights, MN 55120

Phone : 612-452-6650

1100 Lowater Rd.
Cray Research Inc.
Chippewa Falls, Wisconsin 54701

Phone : 715-726-1211

Vector Register Parallel Shared Memory Architecture

This is a 4-processor (quadrant) vector machine with pipelining but no chaining.
There are more segments in the pipes than in the other CRAYs.
Multitasking will be compatible with the X-MP.

The system has a 4-nsec clock cycle time.

Memory is 256 M words of 256 K DRAM in 128 banks. The bank busy time is 57 clocks, and the scalar memory access time is 59 clocks.
Local memory is 16 Kwords, 4 clocks from local memory to vector registers.
Vector references from local memory must be with unit stride.
There are 8 vector registers each with 64 elements.

Overheads for vector operations are large:
63 cycles for vector load
22 cycles for vector multiply
22 cycles for vector add
63 cycles for vector store

The machine is liquid cooled using inert hydrocarbon.

Software:
UNIX-based OS

C compiler
CFT2 (Fortran compiler)

Performance: Max. quoted at 1-2 GFLOPS, but the single pipe to memory and slower access (for vectors) coupled with the lack of chaining means that performances equivalent to X-MP might be possible. Principal advantage of the CRAY-2 is its large memory.

CRAY-3

Cray Research Inc.
1440 Northland Drive
Mendota Heights, MN 55120

612-452-6650

1100 Lowater Rd.
Cray Research Inc.
Chippewa Falls, Wisconsin 54701
715-726-1211

Vector Parallel Architecture

The machine is essentially a GaAs version of the CRAY-2 being developed by a team under Seymour Cray at Chippewa Falls.

Architecture:

16 processors
2-nsec cycle time
4 logical functions/clock period
Memory twice as fast as CRAY-2.
Speed about 8 times CRAY-2.

CRAY-2 imbalance removed by increasing scalar speed to four times that of a CRAY-2 on each processor so, 12x scalar. Aim is 100 times a CRAY-1.

Boards reduced from the 4 x 8 x 1 of the CRAY-2 to 1 x 1 x .1.
Only 1 cu ft in size, with power dissipation of 180 kW as in CRAY-2.
Power supplies take 10 cu ft and liquid coolant 100 cu ft.

Status: 1988 production version; 1990 sales

CRAY X-MP

Cray Research Inc.
1440 Northland Drive
Mendota Heights, MN 55120
612-452-6650

Steve Chen
Chris Hsiung
1100 Lowater Rd.
Cray Research Inc.
Chippewa Falls, Wisconsin 54701
715-726-1211

Vector Register Parallel Shared Memory Architecture

This is a multiprocessor pipelined vector machine. It has the same architecture as the CRAY-1. The major difference is that there are now three paths from memory to the vector registers, and the clock cycle time is now 9.5 ns.

The current machines come with 1, 2, or 4 processors. Gather/scatter hardware is available on the 2- or 4-processor version of the machine. The gather/scatter can can be chained to load/store operation. Users can control both processors through calls in Fortran. The processors share memory.

Other features:
Memory up to 16 M (64-bit) words
X-MP-2 - MOS. (Bank busy time is 76 ns and a memory access time of 17 clocks.)
X-MP-4 ECL. (Bank busy time on the ECL machine is 38 ns and a memory access time of 14 clocks.)
ECL logic with .35-.5 ns gate delay and 16 gates/chip.
Main memory - ECL 4K RAMs with 25-ns access time.
(Interleaving to 64 banks is possible.)

High-speed connection at 1024 Mbytes/sec per channel (max. 2) to a CRAY SSD. The SSD comes in various sizes up to 128 M word of secondary MOS memory. The memory is the same as is used in the CRAY-2. Data transfer to high speed (1200 Mbyte) DD-49 disk takes 10 Mbytes/sec.

Configuration: There are many possible front ends including IBM, CDC, VAX, and Apollo.

Performance: Max. per processor is 210 MFLOPS.

Status: Announced in August 1982, first system delivered in June 1983.

Culler Systems (formally Culler Harris Inc., or CHI)

Culler Systems Inc.
100 Burns Place
Goleta, California 93117
805-683-5631

G. J. Culler

Scalar Pipeline (Array) Architecture

Software: UNIX

CDC CYBER 205

ETA Systems, Incorporated
1450 Energy Park Drive
St. Paul, MN 55108

612/642-3400

Charles D Swanson - Account Support

Vector Architecture

Architecture:
- ECL/LSI logic (168 gates/chip)
- Sequential and parallel processing on single bits, 8-bit bytes and 32- or 64-bit floating-point operands
- 20-nanosecond cycle time

Scalar Unit
- Segmented functional units
- 64-word instruction stack
- 256 word high-speed register file

Vector Unit
- 1, 2, or 4 segmented vector pipelines
- memory-to-memory data streaming
- maximum vector length of 65,536 words
- gather/scatter instructions
- up to 800 million 32-bit floating-point operations/second

Memory
- MOS semiconductor memory
- Memory size: 1, 2, 4, 8 or 16 million 64-bit words
- Virtual memory accessing mechanism with multiple, concurrently usable page sizes
- SECDED on each 32-bit half word
- 48-bit address (address space of 4 trillion words per user)
- 80 nanosecond memory bank cycle time
- Memory bandwidth: 25.6 or 51.2 Gigabits/second

I/O

Eight I/O ports, 32-bits in width, expandable to 16
200 M bits/second for each port
Maximum I/O port bandwidth of 3200 M bits/sec

Miscellaneous

Cooling: freon
Dimensions: floor area (four pipe model) 23 ft x 19 ft
"footprint" (with I/O system) 105 sq ft

Software:

Virtual operating system
Batch and interactive access
FORTRAN compiler
- ANSI 77 with vector extensions
- 32-bit half-precision data type
- Special calls to machine instructions
- Automatic vectorization
- Scalar optimization utilizing large register file

Utilities
- Interactive symbolic debugger
- Source code maintenance
- Object code maintenance

Performance:

Linked triad performance on long vectors approaches asymptotic speed of machine.

Performance can be severely degraded at short vector lengths (that is, the typical $n_{1/2}$ is around 100) and if vector is not held contiguously. For this reason most tuned software employs long, contiguously held vectors.

CYBERPLUS

Control Data Corporation
CYBERPLUS Marketing
P.O. Box 0
HQS09B
Minneapolis, MN 55440

800-828-8001 ext 88

Pete Zidek
612-853-5445

Ring Bus Architecture

This is a multiple parallel processor system. It grew from the Flexible Project and the subsequent Advanced Flexible Processor Project (AFP), used in military applications since 1976. The machine is based on ring technology with an 800 Megabits/second transfer rate with a read and a write possible between processors at this sustained rate.

There are two CYBERPLUS processor models: 16-bit integer and 32- and 64-bit floating point. The integer processor has 15 independent functional units capable of 8-, 16- and 32-bit working; each processor has a 20-nsec cycle time. The floating point processor is an extension of the integer one through the addition of three floating point functional units capable of 32- and 64-bit precision, with rated maximum performance of 65 MFLOPS (103 in 32-bit mode).

Each processor contains 2048 Kbytes of memory which can be expanded to 4096 KBytes. A crossbar architecture allows the output of one functional unit to go to any or all other functional units in one machine cycle and permits all functional units to fire every cycle. There are 15 independent functional units:

- 1 program unit
- 9 I/O units including 4 read/write 16-bit memory units
- 2 read/write 64-bit memory units, 2 ring port I/O units,
- 5 integer/Boolean units (2 add/subtract, 1 multiply, and 2 shift Boolean)

Floating point: 1 add/subtract, 1 multiply, 1 divide/square root connected by an

additional crossbar. Floating-point units can run simultaneously with fixed-point ones.

Each instruction can initiate multiple functional units.

Configuration:

Up to 16 rings can be connected to a CYBER 800 computer (each connected through a channel ring port) with up to 16 CYBERPLUS processors per ring. Within this ring all processors can operate autonomously and may execute each clock cycle. Processor Memory Interface allows direct reading and writing of the memory of any processor by another processor on the ring every machine cycle. Central Memory Interface (CMI) for transfer of data to host. The central memory ring is 64 bits wide with an 80 nanosecond cycle time, and this provides a direct transfer of 64 bits between the CYBER and a Cyberplus processor. Data transfers are controlled by the system ring and will be direct memory-memory transfers with the HPM memory on the CYBERPLUS processors. There are two rings connecting the processors: the system ring and the application ring. The ring packet has 13 bits of control information and 16 bits of data. A function code in the ring packet can determine whether access to other memories (one or several) is direct or indirect, the latter requiring the acceptance by the target processor.

There are three distinct memory systems:

1. 4K 16-bit data memory: 4 independent bi-polar data memories with a one-cycle read/write.
2. 256K 64-bit high-performance data memory: 4 banks with 4-cycle memory access, expandable to 512K 64-bit words with 8 banks.
3. Program Instruction Memory with 4096 200-bit words. Each machine cycle, the instruction memory fetches and initiates the execution of one or all of the parallel functional units. When the floating point option is in use, the size of these memory words increases to 240 words.

The host CDC 170 Series 800 (under NOS 2) loads code into the processors, transmits data from host to processors, and starts and stops processor's task. Software includes a cross assembler (MICA), a CYBERPLUS instructor load simulator (ECHOS), and an ANSI 77 Fortran cross-compiler.

64-bit floating point is 14 decimal accurate with a range of 10^{-293} to 10^{+322}.

32-bit is 7 decimal accurate with range 10^{-39} to 10^{+37}.

Water cooled

Performance: Claimed performance of 64 CYBERPLUS systems linked to a single Control Data 170 Series 800 is 16 billion calculations per second on signal data applications. Change detection algorithm for image processing is about 100 times faster than on a CDC 7600.

Software: Floating point hardware and software delivered in first quarter 1985. Fortran compiler available for research activities fourth quarter 1984 and released April 1985.

Cost: Entry-level CYBERPLUS base processor is priced at $735,000, which includes a 16-bit integer unit and 2.048 Mbytes of memory. With all available options the price is $1.6 million.

Status: Announced formally on October 4, 1983; deliveries started in the first quarter of 1985.

ICL DAP

S. MacQueen/I. Merry
ICL Defence Systems
Lovelace Road
Bracknell RG12 4SN
England

Professor Dennis Parkinson
DAP Support Unit
QMC
Mile End Road
London
England

Bit Parallel Architecture

Configuration: This is an SIMD lockstep machine which operates on multiple data one bit at a time. It has variable-length arithmetic. Configuration is as a grid of processing elements with nearest neighbor connections. There are also row and column data highways (not present on the ILLIAC IV) so that broadcasts can be used to sum efficiently the entries of an array or to find the maximum entry, for example. The other main advantage over the ILLIAC IV lies in the far greater memory for each processing element and the greater reliability of the components.

Three versions of the machine have been produced to date. The first, the prototype 32 x 32 machine, was followed by a larger 64 x 64 version which had an ICL 2900 host. The DAP was configured as one of the host's store modules. This resulted in no communication costs between the two machines when a common data to memory mapping format was used. The standard machine had 2 Megabytes of store, but the QMC (Queen Mary College) machine was later upgraded to 8 Megabytes (i.e., it can be visualized as a cube of dimensions 64 x 64 x 2048 bytes). Six of these machines are in use.

The third version of the machine has returned to the 32 x 32 array size, and has 1 or 2 Megabytes of store. The machine is approximately two orders of magnitude smaller, (it now fits under a desk) and can run without a host. The only architectural change has been the provision of a 40 Megabyte/sec I/O subsystem to permit real time processing. The instruction cycle time has also been reduced from 200 to 150 nsec.

Software: The development environment (cross-compilers and run time debugging aids) are supplied by the ICL Perq running under UNIX. The DAP is linked as a peripheral via a 1.5 Megabyte/sec parallel interface.

Language: The principal programming language used is DAP Fortran, an augmented Fortran that includes most of the array features proposed for Fortran 8X.

Applications: Some of its main applications are in lattice gauge theory and molecular dynamics. It is particularly powerful on the Ising model because of its bit arithmetic. It is also used in many Monte-Carlo calculations and in image processing where the major problem is in data movement rather than processing speed. For some specialized applications, the DAP will outperform a CRAY-1. The new mini DAP has also been used to implement a high-performance military radar system.

Cost: The mini DAP is currently priced at around $300,000 including the Perq and development software.

Status: Work has already begun on a new machine that will use VLSI to achieve further improvements in integration levels and heat dissipation, with a dramatically improved arithmetic performance.

Elxsi System 6400

Elxsi
2334 Lundy Place
San Jose, CA 95131

408-942-1111

Jeff Oromaner - Marketing
Len Shar - Research

Parallel Processor/Bus Architecture

This machine uses ECL technology high-density LSI components. The system can be used as a multiprocessor for multitasking of a single Fortran program, or as a loosely coupled architecture with no parallel processing capability executing independent programs or processes, or both ways.

The system can be configured with 192 Mbytes of memory and many disk drives (474 Mbytes each). Up to 12 processors can be configured with this machine, with 16 Kbytes of cache on each processor.

Global memory architecture is via a fast bus. The bus is 64-bit wide channel providing a gross bandwidth of 320 Mbytes per second, giving a transfer rate 160-213 Mbytes/second. All major components are connected to the bus. Up to 192 Mbytes of MOS memory is available (4 Gbytes virtual).

Other features:

Each CPU 3 boards, each rated at 4 MIPS.
64-bit wide data paths.
50-nsec cycle time.
16-Kbyte, 2-way set associative cache (100-nsec access time).
16 sets of 64-bit general-purpose registers.
IEEE floating point arithmetic.

Software: The operating system, called EMBOS, is a message-based OS. There is also Elxsi's version of UNIX, a port of AT&T System V.2.

Size: The 5-CPU system fits in a single cabinet, 32 in. deep by 59 in. wide.

Languages: Fortran 77, Pascal, COBOL 74, C, MAINSAIL

Cost: A single-processor system is in the range of $400,000.

Encore Multimax

Encore Computer Corp
257 Cedar Hill St
Marlboro, Mass. 01752

617-460-0500

Julius Marcus - VP of Marketing

Parallel/Bus Multiprocessor Architecture

Architecture:

National Semiconductor 32032 chip set running at 10 MHz.
32-Kbyte write-through cache per processor pair.
Processors connected via a fast, 64-bit wide bus
with data throughput rate of 100 Mbytes/sec.
Address space of 4 Gbytes
Main memory 32 Mbytes of RAM in 4 independent banks,
in increments of 4 Mbytes.

Configuration:

Terminal and unit record I/O connected via Annex 16 line terminal
concentrators attached to Ethernet, providing pre-processing.
Is compatable with 19-in. Encore workstation.

Note: The company plans successor chips using best microprocessors,
including RISC architectures.

20 processors maximum configuration.

Ethernet communications using TCP/IP.

Performance: Range quoted from 1.5 MIPS to 15 MIPS by adding processors per module.

Languages: UNIX 4.2 with C, Fortran, and Pascal.

Status: November 1985 with a product

ETA-10

ETA Systems, Incorporated
1450 Energy Park Drive
St. Paul, MN 55108

612/642-3400

Charles D Swanson - Account Support

Vector Parallel Architecture

The ETA-10 is a successor to the CYBER 205, designed to operate at 10 GFLOPS by the end of 1986.

Architecture:

Central Processors

Multiprocessor system with 2, 4, 6, or 8 CPU's
Very high density CMOS circuitry (20,000 gates/chip)
Liquid nitrogen cooling for performance and reliability
CYBER 205 instruction compatibility
Each CPU with a scalar and vector processor, and 4 million words
of local memory

Scalar unit
Independent, segmented functional units
256-word high-speed register file
64-word instruction stack

Vector unit
2 vector pipelines

Memory

Up to 32 million words of CPU memory (4Mw/CPU)
MOS semiconductor Shared Memory using 256K VLSI chips
Shared memory sizes: 64, 128, 192, or 256 million words

1 million word communication buffer for interprocessor communication
Virtual memory addressing
SECDED on each 32-bit half word
48-bit address (address space of 4 trillion words/user)

I/O

Up to 18 400-Mbit/sec Input/Output units for accessing disks, tapes, front-end systems and networks

Miscellaneous

Very low power requirement: 700 watts/CPU (ie about 200 watts per 205 equivalent)
Liquid nitrogen cooling
Compact packaging
High reliability: 100 per cent functional availability

Software:

Virtual operating system
Core operating system for basic processes
User environments for control languages and utilities:
VSOS (CYBER 205 OS - provides CYBER 205 software compatibility)
UNIX

Utilities
Interactive symbolic debugger
Symbolic postmortem dump
Performance analyzer
Source and object code maintenance

Languages:

Fortran
ANSI 77 with vector extensions
32-bit half-precision data type
Special calls to machine instructions
Support for anticipated FORTRAN 8X array notation

Automatic vectorization
Scalar optimization
Multiprocessing library
Pascal
C

Performance: Too early to say. Each processor is claimed to be 3 to 5 times faster than the CYBER 205. Performance will be degraded, however, for noncontiguous vectors and start-up times (%n sub 1/2%) will be high.

Status: Complete system checkout by second quarter of 1986; first shipment by fourth quarter of 1986.

FLEX/32 MultiComputer

Flexible Computer System

Flexible Computer Corporation
1801 Royal Lane
Bldg 8
Dallas, TX 75229

214-869-1234

President/Chairman Larry B. Samartin
President/CEO Dr. M. Nicholas Matelan

William T. Walker
National Manager
Flexible Computer Corporation
5 Great Valley Parkway
Suite 226
Malvern, PA 19355
215-648-3916

Parallel Bus Architecture

This machine is a true 32-bit multicomputer with variable architecture structure and is a MIMD machine. It uses National Semiconductor 32032 chips at 10 MHz, with an independent self-testing system using a Z80 micro. The "local memory" cycle time is 145 nsec. The claimed limit on the number of CPUs is 20480.

Each processor is on one PC with full 32-bit data bus and full 32-bit address capability, with speed capacity of approximatly of 1 MIP using the 32032. Each card has a hardware floating-point processor and hardware memory management and memory protection with a local bus interface and a 32-bit VMEbus I/O interface. Also, each processor board has 1 Mbyte or 4 Mbytes of ECC RAM in addition to cache memory and 128 K of ROM. An optional 1 Mbyte of RAM (later planned to have up to 8 Mbyte) with integral error detection and correction code logic is available. Also, an optional floating point accelerator (1 MFLOP) is available on each processor. The company envisages attaching array processors that are VME compatible such as SKY Warrior.

Other features:

Standard VME bus open architecture supporting Eurocard standard.
Communication rates on local 10 buses 160 Mbit/sec each.
Communicatoin rates on common bus 380 Mbit/sec each.
Time to get on local bus - 1 msec.
Time to do an an arbitrated read/write through high speed (45 nsec) common memory - 170-185 nsec
Direct messaging to another processor's memory via global memory.

Configuration: The machine can have flexible configuration of local (145 nsec) and common memory (45 nsec). Mass memory cards (local memory) contain from 1 to 8 Mbytes RAM connected by local and/or 32-bit VMEbus I/O interface and can be used in any combination or permutation with CPU cards (these memory cards also have a microprocessor for SelfTest diagnostics and fault isolation). The system can be dynamically configured and reconfigured using the SelfTest mechanism.

Software:

A full UNIX System V can run on each processor, with extensions for concurrent processing. FLEX has a 4.2 license. The software license is for 32 users, with optional software license for unlimited users.

FLEX's own multicomputing multitasking operating system (MMOS) is for real-time operating system support providing all the tools for interprocessor communication and signaling, synchronization, event management, etc.

Ethernet-supported TCP/IP

Languages:

Fortran 77 with ISA S61.1 extensions
Ratfor
C
Concurrent C and Fortran by using a preprocessor
Assembly
Ada under development

Base system:

Each cabinet can include up to 20 32-bit processors or 160 Mbytes of memory. There are two computers in two 19-in. standard cabinets:

- one cabinet (the peripheral control cabinet PCC) for the SelfTest System and VME Eurocard card cage (with room for further 19-in. card cages for peripherals)

- the other cabinet (the MultiComputer Cabinet MCC) with a 30-slot card cage partitioned into three 10-slot sections. The backplane contains 2 common buses, 10 local buses, and 20 VMEbus interfaces. The MCC also houses a local bus to common bus interface (common control card) with fair arbitration mechanism up to 9 common access cards with 128 Kbytes to 512 Kbytes of common memory (45 ns) each and a universal card with 128 Kbytes ROM, 1MByte or 4 MBytes of ECC RAM, 1 MIP processor, and VME interface with a separate microprocessor for the SelfTest System.

Cabinet size is 24"x76"x36".

Cost: Price starts at approximately $100,000

$36,000 list price/CPU + 1Mbyte RAM with 128 Kbytes ROM, FPP and MMU.

Floating Point Systems MP32 SERIES   MODEL 3000

MP32 Series, Model 3000,
Floating Point Systems, Inc.

Steve Cannon,
3601 SW Murray Blvd,
Beaverton, OR,
641-3151 x1883

Architecture: MIMD

Basic chip used M68000 (Control Processor), AMD & Weitek Chips (arithmetic processor)
Local, global-shared memory, or both: Both

Connectively (for example, grid, hypercube): Bus

Range of memory sizes available, virtual memory: 1Mword to 7Mword (32-bit)

Floating point unit (IEEE standard?): IEEE standard 32-bit

Configuration
Stand-alone or range of front-ends Front ends: DG MV Series, Perkin-Elmer, Microvax II, VAX

Peripherals: I/O processors

Software: Unix or other? Other

Language available: MAX 68 control language, XPAL assembler

FORTRAN characteristics: N/A
F77
Extensions
Debugging facilities
Vectorizing/parallelizing capabilities: Horizontal microcode synthesis that allows up to 10 operations to execute simultaneously.

Applications:

Run on prototype: Yes, or on front-end simulator

Software available: Math Libraries: Basic math, Signal, Image, & Geophysical

Performance:

Peak: 18 to 54 MFLOPS

Benchmarks on codes and kernels: 2D CFFT 1024 x 1024 pts - 1.89 sec.

Status:

Date of delivery of first machine, beta sites, etc.: Available since 8/85

Expected cost (cost range): $57,500 to $125,000

Proposed market (numbers and class of users): Signal processing, Image processing, and Computational physics

Floating Point Systems
FPS-5000 SERIES

FPS-5000,
Floating Point Systems Inc.

Steve Cannon,
3601 SW Murray Blvd.,
Beaverton, OR,
641-3151, x1883

Architecture: MIMD

Basic chip used: AMD Chips, Weitek Chips on coprocessor
Local, global-shared memory, or both: Both
Connectively (for example, grid, hypercube): Bus
Range of memory sizes available, virtual memory: 256K to 1024K (38-bit words)
Floating point unit (IEEE standard?): 32-bit IEEE (coprocessor)

Configuration:
Stand-alone or range of front-ends: Front ends: VAX; PDP-11; Perkin-Elmer 3200; Gould 32; IBM 4300, 3080, 3090; Prime 750, 9950; Harris 800, HP 1000E
Peripherals: 300MB and 80MB, Disks, I/O processors

Software: UNIX or other? Other

Language available: CP FORTRAN, MAXL control language (FORTRAN-like); APAL and XPAL assemblers

FORTRAN characteristics:
F77 (CPFORTRAN, which is F77 less I/O and character data type support)
Extensions: Calls to coprocessor programs
Debugging facilities: Symbolic debugger
Vectorizing/parallelizing capabilities: Horizontal microcode synthesis that allows up to 10 operators to execute simultaneously

Applications:

Run on prototype: Yes, or run on simulator on front end

Software available: Math Libraries: Basic & advanced math signal and image processing, simulation and geophysical

Performance:

Peak: 8 to 62 MFLOPS

Benchmarks on codes and kernels: 2D convolution 31x31 operations - 33 MFLOPS (FPS-5430)

Status:

Date of delivery of first machine, beta sites, etc.: Oct. 1983

Expected cost (cost range): $45,000 to $99,000 for 256Kword system + standard software

Proposed market (numbers and class of users): 350+ units per year in signal processing, image processing, geophysical analysis, computational physics, and real-time simulation

Floating Point Systems
FPS-164/MAX

FPS-164/MAX,
Floating Point Systems Inc.

Dave Vickers (Technical),
Mike Saunders (Sales)
3601 SW Murray Blvd.,
Beaverton, OR,
641-3151

Pipeline scalar processor with attached processor

Architecture:
- Basic chip used: Proprietary (CPU), Weitek Chips (MAX)
- Local, global-shared memory, or both: Both
- Connectively (for example, grid, hypercube): Bus
- Range of memory sizes available, virtual memory: .5Mwords to 15Mwords (64-bit words) or 4Mbytes to 120Mbytes
- Floating point unit (IEEE standard?): IEEE Standard compatibility

Configuration:
- Stand-alone or range of front-ends: Front end connection to IBM 4300, 308x, 303x, 309x under MVS, MVS/XA, VM/CMS; DEC VAX under VMS; Sperry 1100 Series; Apollo Domain
- Peripherals: FD64 Disk subsystem (1-6 controllers, 4-24 drives), 680MB to 16.2GB

Software: UNIX or other? System Job Executive

Language available: FORTRAN, ASSEMBLY

FORTRAN characteristics:
- F77 ANSI '77 optimizing compiler, 5 levels of optimization
- Extensions: DOE Extensions for asynchronous I/O
- Debugging facilities: Symbolic debugger
- Vectorizing/parallelizing capabilities: Takes advantage of architecture through horizontal micro-coding allowing 10 different operations to occur

in 8 separate functional units per machine/cycle. The matrix algebra accelerator (MAX) modules allow up to 15 concurrent vector operations at any one time.

Applications:

Run on prototype.

Software available: Math Library routines (500+), Fast Matrix Solution Library (FMSLIB) over 40 third party software packages available.

Performance:

Peak: 33-341 MFLOPS

Benchmarks on codes and kernels: 1000 x 1000 Matrix multiply - 66 seconds with 1 MAX module; - 10 seconds with 15 MAX modules

Status:

Date of delivery of first machine, beta sites, etc.: Available since 4/1/85

Expected cost (cost range): $435,000 to $1,900,000

Proposed market (numbers and class of users): Computational Chemistry/Physics, Electronic Circuit Design, Oil Reservoir Simulation, Structural Analysis

Floating Point Systems

FPS-264

FPS-264,
Floating Point Systems Inc.

Dave Vickers (Technical),
Mike Saunders (Sales),
3601 SW Murray Blvd.,
Beaverton, OR,
641-3151

Pipelined Scalar Architecture

Basic chip used: Proprietary ECL implementation
Local, global-shared memory, or both: Both
Connectively (for example, grid, hypercube): Bus
Range of memory sizes available, virtual memory: .5MW to 4.5MW (64-bit words), or 4Mbytes to 36Mbytes
Floating point unit (IEEE standard?): IEEE standard compatibility

Configuration:
Stand-alone or range of front-ends: Front-end connection to IBM 4300, 308x, 303x, 309x under VMS, MVS/XA, VM/CMS; DEC VAX under VMS; Sperry 1100 Series; Apollo Domain
Peripherals: FD64 Disk subsystem (1-6 controllers, 4-24 drives), 680MB to 16.2GB

Software: UNIX or other? System Job Executive

Language available: FORTRAN, ASSEMBLY

FORTRAN characteristics:
F77 ANSI '77 optimizing compiler, 5 levels of optimization
Extensions: DOE Extensions for asynchronous I/O
Debugging facilities: Symbolic debugger
Vectorizing/parallelizing capabilities: Takes advantage of architecture through horizontal micro-coding allowing 10 different operations to occur

in 8 separate functional units per machine/cycle.

Applications:

Run on prototype:

Software available: Math Library routines (500+), Fast Matrix Solution Library (FMSLIB) over 40 third party software packages available.

Performance:

Peak: 38 MFLOPS

Benchmarks on codes and kernels: 1000 x 1000 Matrix multiply 53 seconds

Status:

Date of delivery of first machine, beta sites, etc.: Available since July 1985

Expected cost (cost range): $640,000 to $1,350,000

Proposed market (numbers and class of users): Computational Chemistry/Physicals, Electronic Circuit Design, Oil Reservoir Simulation, Structural Analysis

Floating Point Systems
FPS-364

FPS-364,
Floating Point Systems Inc.

Dave Vickers (Technical),
Mike Saunders (Sales),
3601 SW Murray Blvd.,
Beaverton, OR,
641-3151

Scalar Pipelined Architecture

Basic chip used: Proprietary ECL implementation
Local, global-shared memory, or both: Both
Connectively (for example, grid, hypercube): Bus
Range of memory sizes available, virtual memory: .5MW to 9MW (64-bit words) or 4Mbytes to 72Mbytes
Floating point unit (IEEE standard?): IEEE Standard compatibility

Configuration:
Stand-alone or range of front-ends: Front end connection to IBM 4300, 308x, 303x, 309x under MVS, MVS/XA, VM/CMS; DEC VAX under VMS, Sperry 1100 Series; Apollo Domain
Peripherals: FD64 (same as MAX except capacity) 1-2 controllers, 1-8 disks, 680 MB to 5.44 Gbytes

Software: UNIX or other? System Job Executive

Language available: FORTRAN, ASSEMBLY

FORTRAN characteristics:
F77 ANSI '77 optimizing compiler, 5 levels of optimization
Extensions: DOE Extensions for asynchronous I/O
Debugging facilities: Symbolic debugger
Vectorizing/parallelizing capabilities: Takes advantage of architecture through horizontal micro-coding allowing 10 different operations to occur in 8 separate functional units per machine/cycle.

Applications:
Run on prototype:

Software available: Math Library routines (500+), Fast Matrix Solution Library (FMSLIB) over 40 third-party software packages available.

Performance:
Peak: 11 MFLOPS
Benchmarks on codes and kernels: 1000 x 1000 matrix multiply - 189 seconds

Status:
Date of delivery of first machine, beta sites, etc.: Available since Sept. 1, 1985.
Expected cost (cost range): $298,000 to $950,000
Proposed market (numbers and class of users): Computational Chemistry/Physics, Electronic Circuit Design, Oil Reservoir Simulation, Structural Analysis

Galaxy YH-1

Vector Register Architecture

China has built its first supercomputer, as was revealed by *China Pictorial*. The development of this machine, which has the appearance of a CRAY computer, started in 1978 at the University of Defense, Science and Technology in Changsa.

Performance: The YH-1 (Galaxy), as it is called, can execute 100 million operations per second.

Status: According to *China Pictorial*, the YH-1 was finished two years ahead of schedule and at only one-fifth of the planned budget.

HEP

Denelcor, Inc.
17000 E. Ohio Place
Aurora
Colorado 80017

8-303-337-7900

Dr. Burton Smith - architect

Shared Memory Multiprocessor

The Heterogeneous Element Processor (HEP) is a MIMD machine with two levels of parallelism. Each Process Execution Module (PEM) can run asynchronously, and all can have access to the common storage through a proprietary switch. Although the HEP has been designed for use with up to 16 PEMs, the largest built was a 4-PEM machine. Each PEM is itself an MIMD machine with parallelism achieved through an instruction execution pipeline. Up to 64 user-defined tasks can be executing concurrently, but the length of the pipeline on a 1-PEM machine effectively limits the degree of parallelism to between 8 and 16, depending on memory accesses. The memory accesses are also pipelined. An instruction progresses to the next stage of the pipeline every clock cycle of 100 nsec, although a memory fetch or store can be proceeding simultaneously.

The CPU uses MSI ECL, mostly ECL 10 K with a gate delay of 3 ns, although some critical circuits use ECL 100 K with a .75-nsec gate delay. SECDED memory is used throughout.

Parallelism is obtained in Fortran by explicit task creation (with minimal overhead), and synchronization is by means of asynchronous variables.

Program, constant, register, and data memories all use 64-bit words.

- Program memory size is from 32 Kwords to 1 Mword.
- There are 2048 registers, and the minimum size of the read-only constant memory is 4096 words.
- The data memory is separate from the CPU and can be expanded in 128-Kword increments to a maximum of 1M words (8 Mbytes) per PEM. Memory access time is 50

nsec, and half and quarter word and byte addressing is possible.

Configuration:

The HEP switch that connects memory with CPUs is a flexibly configured, programmable network which uses packet switching techniques to route messages. Each node on the switch network has three full-duplex ports. Arbitration is through a priority system based on longevity. The propagation time through a node is 50 nsec.

Although designed as a stand alone system, it is probably best to front-end the machine with a machine with good interactive capabilities like a VAX.

Software: A version of UNIX III is used as the operating system, although not all utilities are available. The debugging and diagnostic capabilities are poor. Floating point uses IBM-compatible 32- and 64-bit formats. Little software outside of linear algebra kernels is available.

Languages: Fortran 77, C, and Pascal are available in addition to HEP assembler.

Performance: Each PEM is rated at 10 MIPS, and speeds in excess of 7 MFLOPS have been achieved on one PEM for linear algebra kernels coded in HEP assembler language. It is rare to exceed 3 MFLOPS for purely Fortran code on one PEM.

Cost: The cost of a 1-PEM configuration is around $3 million.

Status: There are six systems operational worldwide.

Hitachi S-810

Fred Pazos
Hitachi America Ltd.
950 Elm Ave.
San Bruno, CA 94066

415-872-1902

Vector Register Architecture

The Hitachi comes in two models: the S-810/10 and S-810/20 (not available in the United States, only for the Japanese market).

Hitachi's approach has been to employ independent scalar and vector processors. The S-810/20 relies on their current top-of-the-line mainframe (the M280H) for their scalar processor, with a cycle time of 28 nsec, and runs the complete IBM 370 instruction set. The vector unit was designed with a cycle time of 14 nsec. The main memory capacity of the S-810/20 is 32 megawords.

The model 20 has four floating point add/logical units and eight combination multiply/divide-add units. In addition, there are three load pipes and one load/store pipe to/from memory, each capable of loads/stores at a rate of two word (64 bits) per cycle.

The scalar speed of the Hitachi S-810 may be slower than either the CRAY X-MP or Fujitsu VP-200.

The vector register capacity is 32 registers, each with a fixed length of 256 elements (64 bits). A unique feature of the Hitachi design is that vectors greater than 256 elements are managed automatically by the hardware.

International Parallel Machines Inc. (IP-1)

Robin Chang

International Parallel Machines Inc.
700 Pleasant Street
New Beadford, Massachusetts 02740

617-990-2977

Parallel Architecture

Basic system:
- 9 processors with interconnection switch
- 32-bit addressing
- 4-20 MIPS
- 10 Mbytes main memory
- 150-Mbytex disk
- Over 50 I/O ports
- Timing sharing operating system, up to 8 users.

Can add:
- 1/2-inch magnetic tape drive
- Parallel disk drives
- Character and graphs terminals
- Plotters and printers
- 30 MBytes main memory
- More processors with 60 MIPS
- 20-160 MFLOPS floating point accelerator with optional 20 MFLOPS double precision

Software: UNIX-like operating system.

Application software available: Database management, signal processing, system of equations, computer-aided design of printed circuit boards.

Cost:

9-processor standard configuration with applications software - $49,950.

160 MFLOPS floating point accelerator with application software $74,000.

1/2 inch tape drive - $3,900
CRT terminal w/cable - $400
graphic processor and color terminal - $5,500
36-in. color plotter w/cable - $9,900
150 Mbyte x 8 parallel disks w/software - $49,900

INTEL's Personal Supercomputers (iPSC)

Intel Corporation
15201 NW Greenbriar PW
Beaverton, Oregon 97007

503-629-7629

Justin Rattner
Cleve Moler

Hypercube Architecture

Developed from CALTECH work on Cosmic Cube.
Billed as personal supercomputer.
Controlled by a 286/310 workstation acting as host with
2 Mbytes memory, a 40-Mbyte Winchester disk, and a 320-Kbyte floppy.
This is called a cube manager.
Each node consists of a 80286 and a 80287 processor
(thus, will have IEEE arithmetic 32-, 64-, and 80-bit formats).
512 K RAM per node, 7 bi-directional channels with 10 Mbits/sec per
channel.
Systems range from 32 to 64 to 128 nodes termed the iPSC/d5 through d7.

Channels 10 Megabit/sec Ethernet to nodes

For the 32-node unit:

16 x 16 x 19 inches
footprint 26 x 26 inches
180 lb.

Software: Operating system on 310 Xenix 3.0 with 4.2-compatible interface.

Languages: Fortran and C compilers

Cost and performance of the three machines are as follows:

iPSC/d5	iPSC/d6	iPSC/d7
32 nodes	64 nodes	128 nodes
16 Mbytes	32 Mbytes	64 Mbytes
6 x VAX 11/780	12 x VAX 11/780	24 x VAX 11/780
$150K	$275K	$520K
2 MFLOPS/25 MIPS	4 MFLOPS/50 MIPS	8 MFLOPS/100 MIPS

Status:

iPSC II - 100 MFLOPS by 1987.
iPSC III - 1 GFLOP by 1990 (using 2/4 processors per node with shared memory within the node).

Loral Dataflo

Loral Instrumentation
8401 Aero Drive
San Diego, California 92123

619-560-5888

Parallel Dataflow Architecture

The Loral DATAFLO system is a parallel processor that can be incrementally expanded from approximately 10 processors to approximately 100 processors. Each processor is composed of two National Semiconductor NS32016 microprocessors. One processor is dedicated to collecting data and the other is dedicated to application execution. The application processor has a National Floating Point Unit associated with it. The applications processors each have 128 K of local static RAM that is used for application execution.

In general, communication between processors is via messages (dataflow tokens). Communication is handled on a 32-bit time mutliplex bus. This bus is used to broadcast dataflow tokens that have 16 bits of tag and 16 bits of data. A large dataflow system is composed of multiple chassis, with at most 14 dataflow processors per programmed to pass dataflow tokens between chassis. Since these interfaces pass only those tokens that they are programmed to pass, bus saturation within a chassis is minimized. Shared memory can be added to the system in 2-Mbyte increments by replacing a dataflow processor with a shared memory board. Shared memory can be accessed by any processor in the chassis via a device bus that is separate from the dataflow bus.

A program is composed of two components, a data graph description and a set of graph node implementations written in some standard language like C or Fortran.

The "grain" size for the system is approximately the size of a procedure, around 60 to 100 lines of source code.

MIPS

MIPS
1330 Charleton Rd.
Mountain View, CA 94043

415-960-1200

Scalar RISC Architecture

This is a new organization (approximately 1 year old), with about 60 people, including John Hesseney, John Macerces, and Skip Streater.

Architecture:

Family of CPU boards.
2 - 5 - 8 MIPS (VAX 1.0 MIP)
Custom floating point 3 MFLOPS
IEEE arithmetic

Software: CMOS and UNIX

Status: ship in volume in 7-9 months

Cost: $4,000 for the OEM board

Goodyear MPP

Goodyear Aerospace Corporation
1210 Massillon Road
Akron, Ohio 44315

Ken E. Batcher
216-796-4511

Parallel Architecture

The MPP is the product of research and development designed to evaluate the application of a computer architecture containing thousands of processing elements, all operating concurrently.

The major elements are the array unit, the array control unit, and the staging buffer. The 128x128 processing element has nearest neighbor connection with full-edge closure. The 16,384 processing elements, not including the extra columns for reliability, are simple bit-serial processors, each with a 32 element on chip shift register.

The heart of the array unit is a custom integrated circuit containing eight processing elements. A total of 2112 chips have been combined with commercial memory on control chips to give the capability to perform 400 million floating-point operations per second.

The array control unit contains all the logic to provide a pipeline of commands to the array unit, an I/O controller, and a custom-built 16-bit high-performance microprocessor for program management. The staged buffer is a 16-Mbyte, multidimensional I/O buffer. This unit has the capability necessary to reformat input date into the bit plane format of the MPP I/O system. The staging buffer has an external input rate of 40 Mbytes and an internal transfer rate to and from the array unit of 160 Mbytes in each direction.

Language: Parallel Pascal

Status: The Massively Parallel Processor was delivered to NASA Goddard Space Flight Center in May 1983.

Multiflow

Josh Fisher
Don Eckdahl
Multiflow Computer Inc.
175 N. Main St.
Branford, CT 06405

203-488-6090

VLIW (Very Long Instruction Word) Architecture

Performance: Vector/parallelism capabilities by different techniques

Software: IEEE standard arithmetic and UNIX

Applications: Scientific engineering market

Languages: Fortran, C, Pascal, f77 VAX extensions

Cost: under $1 million

Myrias 4000 System

Myrias Research Corporation
200 - 10328 - 81st Avenue
Edmonton AB T6E 1X2
Canada

403 - 432 1616 Telex 037 - 42759

Martin Walker - R&D Program Manager
UUCP:ihnp4!alberta!myrias!maw

Parallel Architecture, local and hierarchial memory

Each processing element (PE) contains one Motorola MC68000 (10 MHz) and 512 Kbytes of 150-nsec DRAM with high-speed DMA interface to a board level bus. The PE interconnect is in hierarchical clusters: a cluster at one level is composed of subclusters, any pair of which at one level lower is separated by exactly one communications link. The three lowest level clusters contain 1, 8, and 128 PE's, respectively. Physical packaging is in Krate Units (1024 PE's). There are up to 32 I/O paths per Krate, with maximum aggregate data transfer rate of 200 Mbytes/sec.

The system uses virtual memory (32-bit addressing) with the hierarchical PE interconnection providing a distributed cache system. There is no shared central memory. Memory size is 512 Mbytes per Krate.

Totally distributed Kernel performs scheduling of parallel tasks, optimization of machine resources, and management of internal data motion (demand paging); is invisible to the user.

Parallel tasks inherit memory images from their parents, and run in separate memory spaces. Tasks are independent and do not propagate side effects to siblings.

Configuration:

- off-the-shelf components
- two-sided PC boards
- 2 bytes of board
- maintenance by on-site board-swap

- a stand-alone system.
- standard network interface (e.g., VME), or to suit customer.

Arithmetic: 32-, 64-, and 128-bit floating point; 8-, 16-, and 32-bit and arbitrary precision fixed point; IEEE 754 option.

Software:

- UNIX System V and BSD 4.2/Myrias 4000 operating system,
  can reside in different domains
  (collections of PE's dedicated to a single process).
- Upwards compatibility with F77/System V.
- Extensions include PAR DO, recursion, and dynamic array dimensions.
- Will run conforming Fortran programs.
- Will have parallel debugging aids.
- Will have (parallel) mathematical library.

Recursive parallel programming methods allow straightforward implementation of optimal divide and conquer algorithms which can minimize computational complexity.

Languages: Parallel Fortran and Parallel C.

Applications: general-purpose physical modelling (neutron transport, magnetic fusion, rational drug design, chemical engineering, quantum chemistry, aerodynamics, seismic processing and hydrocarbon recovery, geophysics, meteorology, and structural design); data processing (cryptography, image processing and generation, searching and sorting); VLSI design; algebraic manipulation.

Performance: achieved through scalable architecture together with algorithmic reduction of computational complexity.

Status: Prototype expected 1986; commercial deliveries 1987. Cost: Price proportional to performance, expected to be more than $1 million.

NCUBE

Sales Office:
700 E. Baseline Rd., Suite D1
Tempe, AZ 85283

Headquarters:
1815 NW 169th Place
Suite 2030
Beaverton, OR 97006

John Palmer (602)839-7545

Hypercube Architecture

Node Processor
- Custom VLSI
- 11 Interrupt driven DMA channels at 2 Megabytes/sec
- 10 channels for hypercube; 1 for system I/O
- VAX style 32 bit byte addressable architecture
  - 16 general registers (32 bits)
  - complete, orthogonal 2 address instruction set
  - 8,16,32 bit integer and logical operations
  - 32,64 bit IEEE floating point operations
  - 17 addressing modes (eg. autoincr,autodecr,autostride)
- Performance (10 Mhz: approx. VAX 780 with fl.pt. accelerator)
  - 1-2 MIPS (32 bits); .5 MFLOPS (32 bits); .3 MFLOPS (64 bits)
- Memory: 128 Kbytes SECDED about 110 KB available for application

Processor Board (16"x22") contains 64 nodes + 8 MBytes SECDED memory
Host Board (16"x22") contains
- Intel 80286/80287 with 4 Megabytes SECDED memory
- 1 ESMD Disk Interface for up to 4 disks (160,330,500 Megabyte)
- 8 serial RS-232 channels
- 1 parallel Centronics compatible interface
- 3 iSBX interfaces
- 16 Node processors with memory; provide small cube for starter system or 128 DMA channels for larger system

Performance: up to 180 Megabytes/sec bandwidth to hypercube
Graphics Board (16"x22") contains 2Kx1Kx8 frame buffer (768x1024 displayed 60 Hz); color table (16 M color); 180 Mbytes/sec data bandwidth (30 frames/sec); zoom; pan; 16 local NCUBE data nodes; text processor; RS-343 RGB output

Configurations
- NCUBE/ten: 16 to 1024 Nodes; 3 ft cubed; 220 v; 8 KW max; air cooled; 24 slot backplane: 8 for I/O options, 16 for Processor Boards; 160, 330 or 500 Megabyte disk drives and 60 Megabyte cartridge tape
- NCUBE/seven: 16 to 128 Nodes; 15" wide by 2 ft squared; 110v; office environment; 4 slot backplane: 2 for I/O options, 2 for Processor Boards; 160 or 330 MB disk 16 MB tape drive
- NCUBE/four: 4 to 16 Nodes; PC-AT Accelerator (4 Nodes+AT bus interface); up to 4 Boards per AT; for software development

Software
- Axis (Host): Unix style multiuser; distributed file system; EMACS style screen editor with up to 4 windows; cube managed as a device that can be allocated in subcubes
- Vertex (Nodes): Message passing primitives including automatic routing; process debugging support
- Fortran 77 and C are available.

Price: $40K(cabinets+peripherals)+$60K*Host Boards+$100K*Processor Boards (University discount available)
Schedule: Betasites working with I/O systems since February, 1985
Product announcement November 18, 1985, SIAM meeting on Parallel Processing
First complete system shipments in December, 1985

NEC SX-1 and SX-2

Mr. K. Naito
NEC Systems Lab.
1414 Massachusetts Ave.
Boxborough, Massachusetts 01719

617-263-2627

Vector Register Architecture

The SX system has two processors, the Central Processor (CP) and the Arithemetic Processor (AP) sharing the main memory. CP is a front-end mainframe processor where system control programs and user programs run. The AP is a kind of Fortran engine dedicated to user programs executing. Although AX runs in standalone mode, NEC supports its ACOS series mainframes and also IBM mainframe connections.

	SX-1	SX-2
Cycle time	7 ns	6 ns
Number pipes	8 v-pipe	16 v-pipe
Length regs	40K v-reg	80K v-reg

Architecture:

- 16 vector arithmetic pipelines: four identical sets each with an add, multiply, logical, and shift pipe.
- 1000 gate LSIs with 250 picosecond gate delay.
- 1 Kbit bipolar memory with 3.5 nanosecond cache memory access time.
- 256 Megabyte memory (512-way interleaving) with 2 Gigabyte extended memory.
- 64K bit static MOS memory chip with 40 nanosecond access time, giving a memory-to-register rate of 11 Gbytes per second.
- Register-to-register machine with 40 (80 on the SX-2) Kbytes of vector registers.
- register-to-register with far more (and more flexible) vector functional units.

Scalar arithmetic is pipelined (128 scalar registers) and operates in parallel with vector units. The NEC scalar cycle time is faster than the vector, and is segmented and

pipelined to allow more than one pair of operands to progress through the same functional unit concurrently.

Software:
- does not run the IBM instruction set (unlike other Japanese computers)
- Fortran 77 with automatic vectorization. Performance tuning tools available are VECTORIZER/SX and ANALYZER/SX. The compiler vectorizes IF statements, intrinsic functions, and indirect addressing using vector gather and scatter instructions (into temporaries).
- uses its own operating system

Languages: Fortran ALGOL, PL/I, BASIC, Pascal, C, LISP, PROLOG and COBOL. In vector mode only Fortran is supported.

Performance: Maximum rating of the SX-1 is 570 Megaflops and of the SX-2 is 1.3 Gigaflops. Peak performance for the SX-2 will be in the 1.3 Gigaflop range. It appears to be the most powerful of the Japanese supercomputers, and the only one to aggressively address the scalar bottleneck.

Status: expected delivery date in the U.S. is first quarter 1986. The NEC machine is available for benchmarking. NEC has sold the first two of its supercomputers in Japan.

Cost:
SX-1: $10-12 million
SX-2: $14-16 million

PS 2000 (Russian supercomputer)

Vector Architecture

References to the PS 2000 can be found in *Computerworld*, Nov. 26 and Dec. 17, 1984. From these references the design appears based on CDC technology, with a performance of between 400 and 600 Megaflops.

An earlier article in *IEEE Computer*, October 1982, pp. 101-102, by Vladimir Myasnikov, mentioned supercomputer systems called the Elbrus-1 and Elbrus-2 rated at 12 and 125 MFLOPS, respectively but gave no indication of the architecture or any mention of the PS 2000.

The Russians themselves are keeping very quiet about the machine, and it would appear that only their proposed sale to India for $20 million hard currency has made the machine known in the West.

SCS-40

Scientific Computer Systems Corporation
25195 S.W. Parkway Ave.
Wilsonville, OR 97070

503-682-7223

President: Bob Schuhmann
Technical: Carl Haberland

Vector Register Architecture

Architecture:
- register-to-register CRAY-compatible architecture (all CRAY software should run on this machine)
- microcode driven emulator to emulate the CRAY X-MP instruction set.
- 64-bit scientific computer with pipelined, asynchronous functional units.
- multiple pipelined functional units.
- 45-nsec cycle time.
- 5 vector, 1 scalar, and an address calculation can execute concurrently.
- transfer rate from registers to functional units of up to 6 words/ clock cycle (1.07 Gbytes/sec).
- 256-word buffer between memory and instruction decode logic allows execution of one instruction per cycle (two cycles for conditional branch).
- supports flexible hardware chaining of functional units and memory references.

Configuration:
- 8-, 16-, 32-Mbyte field-upgradable memory configurations with 4-16 banks.
- four ports to memory (like the CRAY X-MP, i.e., 2 vector loads and a store can be going on at the same time.)
- designed as to interface to a front end, either VAX 11/780 or VAX 11/750. (Interfaces planned for CRAY X-MP, IBM 4300 series, and NSC hyperchannel.)
- 2-10 programmable I/O channels, each with 16 Kbyte buffer and a transfer rate of 20 Mbyte/sec. Transfer rate of buffers to central memory is 1 word/clock period (178Mbytes/sec).
- DD-550 disk drive holds 550 Mbytes and can sustain read/write data transfer rate of 10 Mbyte/sec with an average access time (seek plus latency) of 24 msec
- maximum of eight drives can be attached to each of the eight optional I/O

channels.

Other features:
- Size: 55 x 55 x 60 inches
- Forced air cooling.
- Power consumption: 208 3-phase 11-16.5 KVA
- Weight: 1 ton

Software:
- software licensing agreement with CRAY.
- multiuser, multiprogramming OS supports interactive job execution.

Languages:
- Fortran 77
  Fortran compilation expected at 20,000 to 40,000 lines per minute.
  Fortran vectorizing compiler.
  Interactive debugger
- Assembler

Performance:
- peak of 44 MFLOPS in 64-bit arithmetic
- LINPACK timings around 1/4 the performance of a single CPU X-MP.
- matrix vector operations (subroutine SMXPY), around 37.6 MFLOPS (simulated).

Status: Prototype available 11/85; first customer shipment 4/86

Cost: Base system $500,000. Market target is to provide a CRAY-compatible general-purpose scientific computer that computes at 1/4 the CRAY X-MP, but has the price of a super-mini and thus the price/performance of a supercomputer.

Sequent Balance 8000

Sequent Computer Systems, Inc.
15450 SW Koll Parkway
Beaverton, Oregon 97006-5903

503-626-5700
800-854-0428
Telex 296559

Casey Powell and Scott Gibson, co-founders.
Technical: David Rodgers and Gary Fielland

Chicago Office
Karl von Spreckelsen
Field Sales Engineer
Forum I Suite 12
1117 South Milwaukee Avenue
Libertyville, IL 60048
312-680-9747

Parallel Bus Architecture

Incorporated in January 1983
(old name of company was Sequel)

Machine has 2-12 NS 32032 processors running at 10 MHz, each with floating point unit, memory management unit, and 8-Kbyte cache sharing a global memory via a 32-bit wide pipelined packet bus supporting multiple, overlapped memory and I/O transactions with a sustained data transfer rate of up to 26.7 Mbyte/sec.

Memory: The machine has up to 28 Mbytes of physical memory, a 4-Mbyte I/O address space, and a 16-Mbyte virtual memory address space for each user process. Memory can be two-way interleaved, and there can be up to 4 memory controllers which each manage 2 to 8 Mbytes using 256K-RAM components. Processor and memory boards can go in any slot on the SB8000 bus.

A Sequent-designed IC chip (SLIC, System Link, and Interrupt Controller) resides on each board to manage interprocessor communication, synchronization, interrupts,

diagnostics, and configuration control. There is an extensive diagnostic subsystem.

Software: The operating system, called DYNIX, is a version of Berkeley 4.2bsd UNIX, enhanced for application-transparent multiprocessing and user-controlled parallel processing. Among the enhancements are a completely reentrant kernel, user-level shared memory, and synchronization services. All processors run a shared copy of the operating system. The configuration is symmetric, and load balancing is automatic and dynamic.

Industry-standard I/O, interfaces:
MULTIBUS - has terminal multiplexor with controllers
Ethernet - at 10 Mbits/sec. Connnection to PC as virtual disk through Ethernet.
SCSI - at 2.5 Mbyte/sec. Offers 5-1/4 in. disk drives (72 Mbytes formatted) and streamer tape drives with adaptor boards for the SCSI bus.

Peripherals include a 1/2" tape drive and a 396-Mbyte disk drive asynchronous

The packaged system includes a 9-slot SB8000 bus backplane and an 8-slot MULTIBUS backplane; can take up to six dual processor boards.

Other features:
Table height packaging.
Dimensions 30.5" x 23.25" x 28.625" (HWD)
SB800 chassis 15.5" x 10.5" x 13.5"
MULTIBUS chassis 14.2" x 6.68" x 8.5"
11 amps max at 60Hz 115VAC.
Maximum configuration dissipates 1500 Watts

Software: supports ARPANET TCP/IP protocols plus all the networking facilities of UNIX 4.2. Support is also available for customer-provided application accelerators.

Languages:
C
Fortran 77
ANSI-standard Pascal
Assembly language

Parallel programming library callable from any language.
Extension to Fortran to allow shared common blocks.

Performance:

Fully populated machine seen as 6 times a VAX 11/780 in power.

Designed as a high throughput system, with support for parallel processing at user level.

Status: Shipments began 12/84, and Sequent currently has manufactured more than 80 systems.

Cost:

$220,000 for the complete machine with all software, 12 processors, 8-Kbyte cache/processor, 12-Mbyte memory, and a Fujitsu Eagle 396-Mbyte disk; $62,000 for a small 2-processor system.

ST-100

Star Technologies Inc.
1700 US Bancorp Tower
111 S.W. 5th Avenue
Portland, Oregon 97204

503-227-2052

Technical: Phil Cannon

Pipeline Floating Point Architecture

The ST-100 is an array processor, designed to attach to a more general-purpose computer or host via bus.

It has four independent programmable processors. A separate processor is dedicated to each of the following functions: external data flow, internal data flow, arithmetic processing, and synchronization. A hierarchical memory system consists of external storage devices, a large main memory, a high-speed random access partitioned data cache, and a universal register set.

The main memory consists a 320 nsec memory, 8-way interleaved, composed of 64K dynamic RAMs with SECDED. It is expandable to 32 Mbytes in increments of 2048 Kbytes. All main memory is byte addressable (address range 4 Gbytes) and can be partitioned and protected at multiples of 16 Kbytes. Memory access time is 40 nsec (per 32-bit word). The random access data cache memory consists of 6 banks of 8192 32-bit words for a total of 192 Kbytes. During each machine cycle, four cache references are permitted: three by the arithmetic processor and one by the storage/move processor. Information flow from host to main memory to cache to functional unit to cache to memory to host.

Other features:
- 40 nsec clock cycle,
- Bipolar VLSI circuits with 1200 gates.
- 32-bit floating point arithmetic, pipelined functional units.
- 2 adders, 2 multipliers, and a 480 nsec divide/square root functional unit.
- Ambient air cooled
- Size 56" x 33" x 67"

A data interchange unit permits one of 16 operands to be selected for each arithmetic input register. During each machine cycle, three cache banks may be referenced, one loop control operation computed, four arithmetic operations started, and a conditional branch executed.

The 25 Mbyte I/O channel supports 7 device adapters; 12.5 Mbyte/sec data transfer rate.

Software:
- Fortran-like control language (APCL)
- Macro assembler
- Simulator/debugger and Linker
- Library Maintenance Program
- Applications Library available.

Primary programming accomplished through microcoding VAST (available on the STAR from Pacific Sierra)

Performance: 100 MFLOPS peak in single-precision (32-bit) arithmetic for convolution and matrix operations.

Cost: $265,000 base price.

Vitesse Electronics

741 Calle Plano
Camarillo, CA 93011

805-388-3700

Creve Maples

Parallel Architecture

Plans are to build a scalar machine with with 1Gbyte memory and a 40-nsec cycle time. The machine will be made of CMOS. It is to support hardware optimization for high run-time performance.

Configuration: First machine is to have up to 8 processors. Connectivity allows for large number of processors, in the thousands. It can be used as a co-processor on a VAX.

Software: 32-64 bit floating point arithmetic supporting the IEEE standard.

Languages: Fortran, Pascal, and C.

Performance: 25 to 150 MFLOPS (uniprocessor range of performance as result of optional hardware boards for each processor).

Status: Started in July 1984, expect to produce a machine by late 1986. It is planned to make a GaAs version in a couple of years.

APPENDIX C

UNIVERSITY/RESEARCH COMPUTERS/ARCHITECTURES

Blue Chip, Larry Snyder, University of Washington
Cedar, David Kuck, up to 32 processors using Alliant, University of Illinois
Chorus, Waterloo
Clementi's 10-FPS-164, IBM
CLIP, University College London
CMU-APS, Carnegie Mellon
Columbia 1 (Nonvon), David Shaw, Columbia University
Columbia 2 (DADO), Columbia University
Columbia 3, Norman Christ, 16 processors, Columbia University
Cosmic Cube, hypercube design, CalTech, Fox/Seitz
Dataflow, John Gurd, Manchester University
Dataflow 1, MIT (Dennis)
Dataflow 2, MIT (Arvind)
FEM, NASA Langley
GF11, IBM
Gigacomputer, Argonne National Laboratory
Marianne/Marisa, ONERA-CERT, France
NEPTUNE, Loughborough, England
PASM, Purdue University
Pringle, Purdue University
RP3, Courant/IBM
SIGMA-1, Japan
S1, L. Wood, LLNL
Transputer, INMOS
Ultra, Courant
ZMOB, University of Maryland